The Dissertation Committee for Pawat Seritrakul Certifies that this is the approved version of the following dissertation:

FUNCTIONAL ANALYSIS OF DNA METHYLATION AND HYDROXYMETHYLATION DURING EYE DEVELOPMENT

Committee:

Steven A. Vokes, Supervisor

Jeffrey M. Gross, Co-Supervisor

Zengjian Jeffrey Chen

David S. Stein

Vishwanath R. Iyer

FUNCTIONAL ANALYSIS OF DNA METHYLATION AND HYDROXYMETHYLATION DURING EYE DEVELOPMENT

by

Pawat Seritrakul

Dissertation

Presented to the Faculty of the Graduate School of

The University of Texas at Austin

in Partial Fulfillment

of the Requirements

for the Degree of

Doctor of Philosophy

The University of Texas at Austin

May 2018

Dedication

To my parents, Dr. Wilailuk and Mr. Suvit Seritrakul.

Acknowledgements

I would like to thank my advisor, Dr. Jeff Gross, for his mentorship and for giving me the opportunity to work in the lab, and my Dissertation Committee members, Dr. Steven A. Vokes, Dr. Z. Jeffrey Chen, Dr. David Stein, and Dr. Vishwanath R. Iyer, for their guidance throughout my graduate career.

I am fortunate to have had a series of great mentors, from my biology teachers at Mahidol Wittayanusorn high school, Dr. Bualuang Faiyue, Dr. Apisit Thipaksorn, and Dr. Orawan Piyaboon, who instilled the scientific curiosity in me, to my undergraduate advisors at Bowdoin College, Dr. Hadley Horch and Dr. William Jackman, who introduced me to developmental genetics and the zebrafish model organism.

I would like to thank Dr. Dennis Kostka and members of the Kostka Laboratory for their bioinformatics guidance. I would also like to thank my classmates and colleagues both at Institute for Cellular and Molecular Biology at the University of Texas at Austin, and Department of Ophthalmology, University of Pittsburgh.

Thank you to all past and present members of the Gross Laboratory for making my lab experience exceptionally awesome. Special thanks to my Gross lab comrades, Nick Hanovice, Krista Angileri, and Natalie Gath, who offered endless intellectual and technical support and countless hours of Dota.

Last but not least, I would like to thank my family and friends on both sides of the Pacific Ocean for always being there for me and for believing in me.

FUNCTIONAL ANALYSIS OF DNA METHYLATION AND HYDROXYMETHYLATION DURING EYE DEVELOPMENT

Pawat Seritrakul, Ph.D.

The University of Texas at Austin, 2018

Supervisor: Steven A. Vokes

Co-Supervisor: Jeffrey M. Gross

DNA methylation is an epigenetic mechanism known to play roles in regulating gene expression in various developmental and disease contexts. However, little is known about its function during eye development. Two types of methylation marks, 5-methylcytosine (5mC) and 5-hydroxymethylcytosine (5hmC), are thought to serve as silencing and activating signals for gene regulation, respectively. *De novo* methyltransferases (dnmt3 family) are responsible for the establishment of 5mC, while cytosine dioxygenases (tet family) convert 5mC into 5hmC, a stable epigenetic mark that can either remain on the genome or undergo subsequent demethylation. Here I performed gene expression and functional tests to elucidate the roles for both of these cytosine-modifying enzyme families during development, with an emphasis on the eye. All dnmt3-family and tet-family genes are expressed tissue-specifically in relevant domains during eye development. Single and double mutants for genes within dnmt3 family develop normally without any overt eye phenotype, indicating that these genes possess redundant functions during eye development. In contrast, in *tet2*$^{-/-}$*;tet3*$^{-/-}$ mutants, retinal neurons are specified but most fail to terminally differentiate. Retinal ganglion

cells lack a proper retino-tectal projection, and photoreceptors fail to generate outer segments. Mechanistically, mosaic analyses revealed a surprising cell non-autonomous requirement for tet activity during retinal neurogenesis. Through a combination of candidate gene analysis, transcriptomics and pharmacological manipulations, I identified candidate cell-extrinsic pathways regulated by tet2 and tet3. Additionally, genome-wide 5mC and 5hmC distribution profiles for retinal progenitor cells (RPCs) and differentiated retinal neurons are still currently unknown. To this end, I performed parallel bisulfite and oxidative bisulfite reactions followed by next-generation sequencing (BS/OXBS-seq) to generate the first nucleotide-resolution combined methylome/hydroxymethylome map of retinal cells during development and correlated these with gene expression. This genome-wide approach revealed expected 5mC/5hmC profiles of candidate retinal developmental genes, and identified several novel, uncharacterized genes with potential roles during RPC differentiation. These genes are candidates for further investigation to determine their functions during retinal neurogenesis. Data presented in this Dissertation uncover the role of DNA methylation and hydroxymethylation during eye development and provide the first epigenomic maps of 5mC/5hmC dynamics during retina formation.

Table of Contents

List of Tables

List of Figures

Chapter 1: Introduction to Eye Development and DNA Methylation

Portions of this Chapter are modified from the following articles, with permission from the authors: 1. Seritrakul, P., Gross, J.M., 2014. Expression of the De Novo DNA Methyltransferases (dnmt3 – dnmt8) During Zebrafish Lens Development. Developmental Dynamics. 243:350–356; 2. Seritrakul, P., Gross, J.M., 2017. Tet-mediated DNA hydroxymethylation regulates retinal neurogenesis by modulating cell-extrinsic signaling pathways. PLOS Genetics. 13(9): e1006987
Seritrakul, P. wrote the manuscripts. Gross, J.M. edited the manuscripts.

Neural Retina	NR
Retinal Progenitor Cell	RPC
Ciliary Marginal Zone	CMZ
Retinal Ganglion Cell	RGC
Photoreceptor	PR
Ganglion Cell Layer	GCL
Inner Nuclear Layer	INL
Outer Nuclear Layer	ONL
Lens Epithelium	LE
Lens Fiber	LF
Transcription Factor	TF
5-methylcytosine	5mC
5-hydroxymethylcytosine	5hmC
Differentially Methylated Region	DMR
Differentially Hydroxymethylated Region	DhMR
Bisulfite Sequencing	BS
Oxidative Bisulfite Sequencing	OXBS

Table 1.1: List of common Abbreviations

1.1 Development and Function of the Neural Retina

The vertebrate neural retina (NR) is a complex structure consisting of six major neuronal types and one glial cell type that work together to perform phototransduction and transmit visual information to the brain. These cells are organized into three histologically distinct retinal layers: the ganglion cell layer (GCL), outer nuclear layer (ONL), and inner nuclear layer (INL). All seven retinal cell types are born in a highly stereotyped chronological order, a process called neurogenesis, with the first-born cell type being retinal ganglion cells (RGC), followed by horizontal cells (HC), cone photoreceptors, amacrine cells (AC), rod photoreceptors, bipolar cells (BP), and lastly Müller Glia cells (MG) [reviewed in (Bassett and Wallace, 2012)]. Functionally, rods and cones (collectively referred to as photoreceptors: PRs) detect and convert photons of light into synaptic signals, which are transmitted through a series of interneurons (HC, BP, AC) to RGCs, which relay this signal, via the optic nerve (ON), to the optic tectum of the brain. Additionally, some vertebrate species, such as frogs and teleost fishes, also possess a unique niche of retinal stem cells (RSCs) located in a region of the peripheral retina called the ciliary marginal zone (CMZ). This structure persists throughout the life of the animal and continuously adds more differentiated cells as the eye continues to grow (Raymond et al., 2006).

1.1.1 Early retinal development

The onset of NR development in vertebrates begins as the optic vesicles (OV) evaginate as pouches from each side of the neural plate, then undergo elongation, forming wing-like structures attached to the brain via the optic stalk (Walls, 1963). Each OV then invaginates to surround the emerging lens originating from the surface

ectoderm, thus completing the morphogenesis of the optic cup (OC), the basic structure of the embryonic eye (Fig 1.1) (Schmitt and Dowling, 1994). During this process, the optic vesicle, composed of the lateral layer (LL) and medial later (ML), undergoes morphogenetic movement to form the presumptive NR and retinal pigmented epithelium (RPE) (Li et al., 2000). Patterning of the early eye is established and maintained by key transcription factors (TFs) such as Vsx2 (formerly Chx10) and Mitfa, for NR and RPE, respectively (Horsford et al., 2004). Cell tracking data showed that the process of OC morphogenesis involves an intricate and well-orchestrated cell movement between RPE and NR surrounding the lens (Kwan et al., 2012). In zebrafish, formation of the OC is completed by 24-hours post-fertilization (hpf) and the first RGCs become morphologically identifiable by 32 hpf (Hu and Easter, 1999; Schmitt and Dowling, 1999).

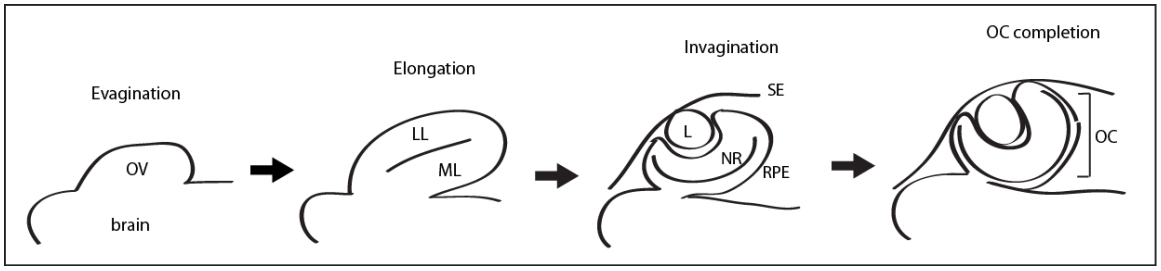

Figure 1.1: Schematic of optic cup morphogenesis during early eye development. OV, optic vesicle; LL, lateral layer; ML, medial layer; L, lens; NR, neural retina; RPE, retinal pigmented epithelium; SE, surface ectoderm; OC, optic cup.

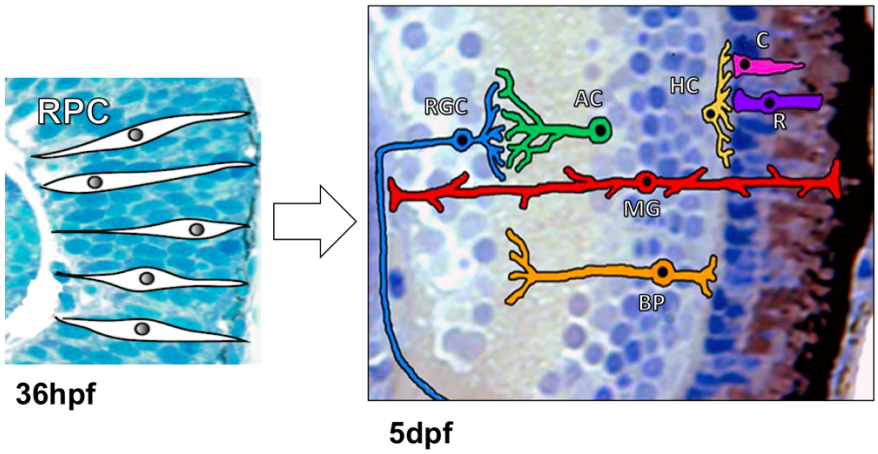

Figure 1.2: Retina progenitor cells (RPCs) give rise to all differentiated retinal neurons and Müller Glial cells. Image modified from Bibliowicz et al., 2011.

1.1.2 Retinal cell type determination

During neurogenesis, all seven retinal cell types arise from a common pool of retinal progenitor cells (RPCs), a population of seemingly indistinguishable multipotent cells organized as a pseudostratified epithelial layer (Fig 1.2) [reviewed in (Centanin and Wittbrodt, 2014)]. The first evidence that RPCs are capable of giving rise to all differentiated cell types in the mature retina comes from early lineage tracing experiments, where any single labeled RPCs in the OV can proliferate and differentiate into clones composed of all seven cell types (Wetts and Fraser, 1988). How a homogeneous population of RPCs undergoes transformation into diverse cell types that together build such an elegant and complex structure as the retina has been a subject of intense investigation. To explain this phenomenon, two distinct but not mutually exclusive models have been proposed.

First, the Competency Model builds upon the observation that retinal cells are born in a histologically and chronologically stereotyped fashion, and posits that RPCs

gradually become more restricted in their differentiation potential over time (Belliveau and Cepko, 1999; Wong and Rapaport, 2009). In other words, the earliest RPCs are competent to generate all seven cell types, while later RPCs can only generate a few late-born cell types as the developing retina reaches the final phases of neurogenesis, becoming fully mature and functional. The competency model is consistent with the idea that RPCs are pre-programmed by an intrinsic circuitry that operates internally on a cell-autonomous schedule and are largely uninfluenced by any external environmental cues. This is supported by early mixed-age cell culture experiments where undifferentiated RPCs progress down the predetermined path and differentiate into all seven cell types even when co-cultured among already differentiated cells from a mature retina, relatively unaffected by their extracellular environment (Cayouette et al., 2003; Rapaport et al., 2001; Watanabe and Raff, 1990). The competency model is reminiscent of the neuroblast developmental program in Drosophila, where Hunchback triggers a cascade of TFs to generate neurons in a temporally restricted lineage (Isshiki et al., 2001). In vertebrates, Ikaros is a Hunchback ortholog that performs a similar role in early-born retinal cell differentiation (Elliott et al., 2008) suggesting that this intrinsic transcriptional cascade is a highly conserved mechanism for neuronal specification.

Alternatively, the Stochastic Model was more recently developed based on observations from lineage tracing experiments that show individual RPCs differentiate into clones of variable sizes and cell types (Gomes et al., 2011; He et al., 2012). However, collectively as a population, the final number and composition of differentiated cells among these clones can be explained by a mathematical model based purely on probability and stochasticity (Slater et al., 2009). For example, if the final retina is composed of 70% PRs and 5% RGCs, the probability of generating PRs and RGCs from

any given RPC is 70% and 5%, respectively. Additionally, these lineage tracing experiments detected some RPCs giving rise to neurons in a reverse order (i.e. late-born cells before early-born cells), reinforcing the idea that fate commitment, though set by cell-intrinsic properties, possesses a certain degree of stochasticity that is somehow regulated and not random. Finally, the model has been directly tested and verified experimentally by knocking down individual key TFs (*atoh7, ptf1a, vsx1*) followed by blastomere transplantation and a careful clonal analysis of RPC daughters (Boije et al., 2015). Therefore, the invariant retinal cellular composition may simply be the result of independent probabilistic and combinatorial transcription of a few key TFs at the core of retinal cell fate determination network. Nonetheless, relatively little is known about the factors that regulate the activity of these TFs in such a way that the correct cell types are born in the right number at the right time.

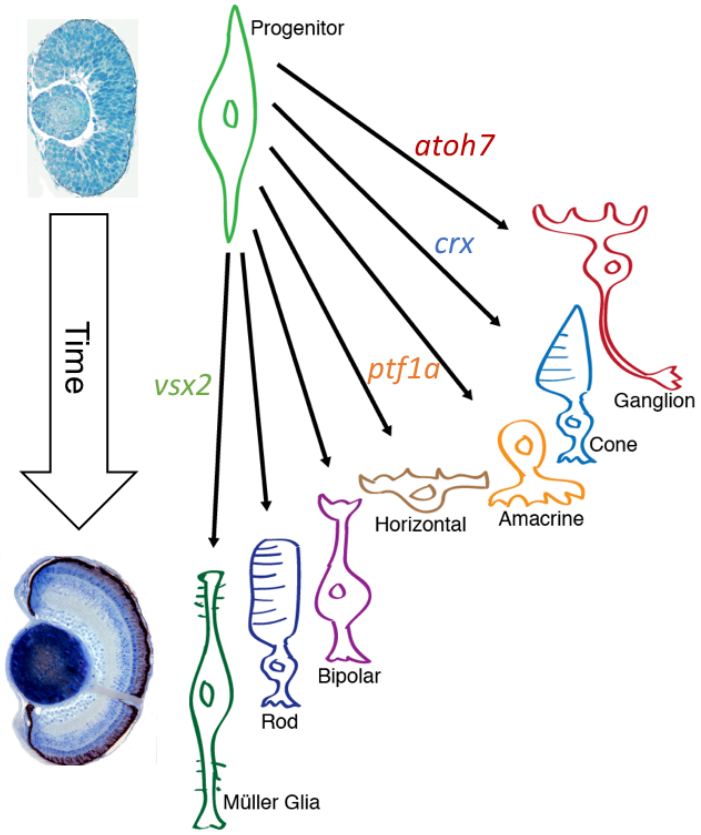

Figure 1.3: Retinal neurons and Müller glial cells undergo differentiation in a chronologically stereotyped fashion from retinal progenitor cells (RPCs) via a cascade of transcriptional activity.

1.1.3 Intrinsic and extrinsic factors regulating retinal development

A multitude of intrinsic and extrinsic signals, are known to regulate retinal development from the establishment and maintenance of the RPC pool to the specification and terminal differentiation of neuronal and glial cells (Fig 1.3). First, the existing RPC population expands in order to reach sufficient numbers before beginning to differentiate. Numerous TFs are known to intrinsically regulate RPC multipotency and proliferation capacity, the most well-characterized being Pax6 (Marquardt et al., 2001),

Vsx2 (Vitorino et al., 2009), and Sox2 (Matsushima et al., 2011). Loss of Pax6 in RPCs results in a hypocellular retina that lacked virtually all subsequent differentiation events (Klimova and Kozmik, 2014), while loss of Sox2 leads to decreased RPC neurogenic competency and conversion of RPCs to undifferentiated ciliary epithelium (Matsushima et al., 2011). In addition to their roles in RPCs, Pax6 and Vsx2 also have neurogenic roles later in differentiating retinal neurons, where Pax6 is required for RGC and AC differentiation (Oron-Karni et al., 2008) and Vsx2 is required for BP and MG differentiation (Vitorino et al., 2009).

Downstream of RPC-specific genes are a cascade of TFs required for specification of each retinal cell type. For example, the proneural TF, Atoh7 (formerly Ath5), is expressed during the last division of RPCs and is required for RGC specification; inactivation of Atoh7 results in a near complete loss of RGCs and an increase of later born cell types (Kay et al., 2001). Atoh7, in turn, activates other TFs such as Pou4f1-3/Brn3a-c (Liu et al., 2000) and Isl2 (Pan et al., 2008), which are required for expression of subsequent RGC terminal differentiation genes including components of neurotransmission, axonal growth, and pathfinding. Similar scenarios occur in other retinal cells where a combination of key TFs specific to each cell type activate a cascade of downstream TFs and specialized cell-specific molecules. These include Crx for rod and cone photoreceptors (Furukawa et al., 1997), Ptf1a for ACs (Jusuf and Harris, 2009), and NeuroD4 for HCs and BPs (Wang et al., 2003).

In addition to the intrinsic signals above, many extrinsic signaling molecules also act upon RPCs during neurogenesis. For example, the Wnt and Notch pathways are both required for inhibiting RPC differentiation and maintaining RPCs in a proliferative state. Inactivating Notch signaling triggers early RPC exit from the cell cycle, and premature differentiation into the first retinal cell fate available, the RGCs, (Bernardos et al., 2005;

Perron and Harris, 2000). Conversely, over-activation of Wnt components results in enhanced RPC proliferation and down-regulation of proneural bHLH TFs, thus preventing exit from the cell cycle and subsequent differentiation (Kubo et al., 2005). Aside from Notch and Wnt, many other extracellular signaling pathways such as Hedgehog, FGF, VEGF, BMP, and neurotransmitters such as Ach, GABA, glutamate, and dopamine are also involved in retinal neurogenesis [reviewed in (Agathocleous and Harris, 2009)]. Despite the breadth of knowledge about the roles of intrinsic and extrinsic factors during retinal development, relatively little is known about how these factors are regulated initially. Increasing amounts of data are pointing to epigenetic mechanisms, including histone modification (via enzymes such as Hdac1 (Yamaguchi et al., 2005), Brg1 (Das et al., 2007)), and DNA methylation and hydroxymethylation (Merbs et al., 2012; Rhee et al., 2012) as potential regulators of these intrinsic and extrinsic factors, thus adding an extra layer of control to an already complex system.

1.2 Development and Function of the Lens

The vertebrate lens is composed of only two cell types: lens epithelial (LE) and lens fiber (LF) cells. During development and continuously throughout life, proliferative LE cells in the anterior layer of the lens undergo differentiation within the transition zone (TZ) along the lens equator, giving rise to elongated LFs, which are then arranged in an orderly fashion into the outermost lens cortex and comprise the bulk of the lens material (Fig 1.4) (Greiling and Clark, 2009; Greiling et al., 2010). LFs are transparent, highly organized, and organelle-free. To obtain these unique biophysical properties, LFs undergo drastic morphological changes during differentiation that include a massive elongation and complete elimination of all light-scattering organelles, including the

nucleus (Bassnett et al., 2011; Wride, 2011). Failure of LF cells to achieve their final shape and organization have been shown to cause congenital cataracts in human and other model organisms (Hejtmancik, 2009; Santana and Waiswol, 2011).

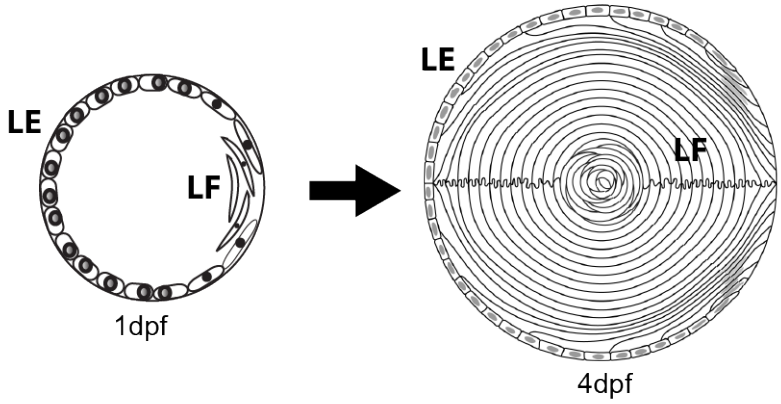

Figure 1.4: Lens epithelial (LE) cells continuously give rise to lens fiber (LF) cells, which are packed into the bulk of the lens cortex.

1.3 DNA METHYLATION PATHWAY

DNA methylation is an epigenetic mechanism by which methyl groups are added to the 5th carbon of cytosine residues (5mC), predominantly at CpG dinucleotides (Jones, 2012; Suzuki and Bird, 2008). Genomic regions rich in CpG dinucleotids (CpG islands) are often associated with gene promoters, and their 5mC methylation statuses are often inversely correlated with gene expression levels (Deaton and Bird, 2011; Illingworth and Bird, 2009; Meissner et al., 2008; VanderKraats et al., 2013; Yamaguchi et al., 2012). Hypermethylation within promoters and enhancers is associated with reduced gene transcription (Wu et al., 2010), while gene body methylation directly correlates with expression (Jin et al., 2012). Indeed, DNA methylation is critical for silencing of

imprinted genes and transposons (Bird, 2002; Li et al., 1993; Walsh et al., 1998). Subsets of genes are differentially methylated according to tissue and cell type, and DNA methylation is thought to be a mechanism whereby cell type-specific expression patterns are set during terminal differentiation (Ehrlich, 2003; Illingworth and Bird, 2009), and by which some somatic progenitor cell populations are maintained (Bröske et al., 2009; Meissner et al., 2008; Sen et al., 2010; Trowbridge et al., 2009).

Three main biochemical events orchestrate DNA methylation (Fig 1.5-6). First, *de novo* methylation, mediated by DNA methyltransferase (Dnmt3-family) proteins, functions to methylate regions of hypomethylated DNA and is required for tissue-specific differentiation during development (Feng et al., 2010; Hu et al., 2012; Illingworth and Bird, 2009; Merbs et al., 2012). Second, maintenance methylation, mediated by Dnmt1, copies the methylation pattern from existing DNA strands on to nascent daughter strands during DNA replication, a process that is important for maintaining the identities of actively proliferating cell populations (Sen et al., 2010; Tittle et al., 2011). Third, demethylation is the mechanism by which 5mC is removed from the genome. Far less is known about the DNA demethylation process but several biochemical pathways have been proposed to be involved and these include: replication-dependent passive dilution, direct base excision by the DNA repair machinery, and active enzymatic conversion of 5mC [reviewed in (Wu and Zhang, 2014)]. Of these pathways, most evidence supports the latter and a role for members of the Ten-eleven Translocation (Tet) family of methylcytosine dioxygenases. These enzymes mediate the conversion of 5mC to 5-hydroxymethylcytosine (5hmC), which can then be converted to non-methylated cytosine (Ito et al., 2010; Kohli and Zhang, 2013; Pastor et al., 2013).

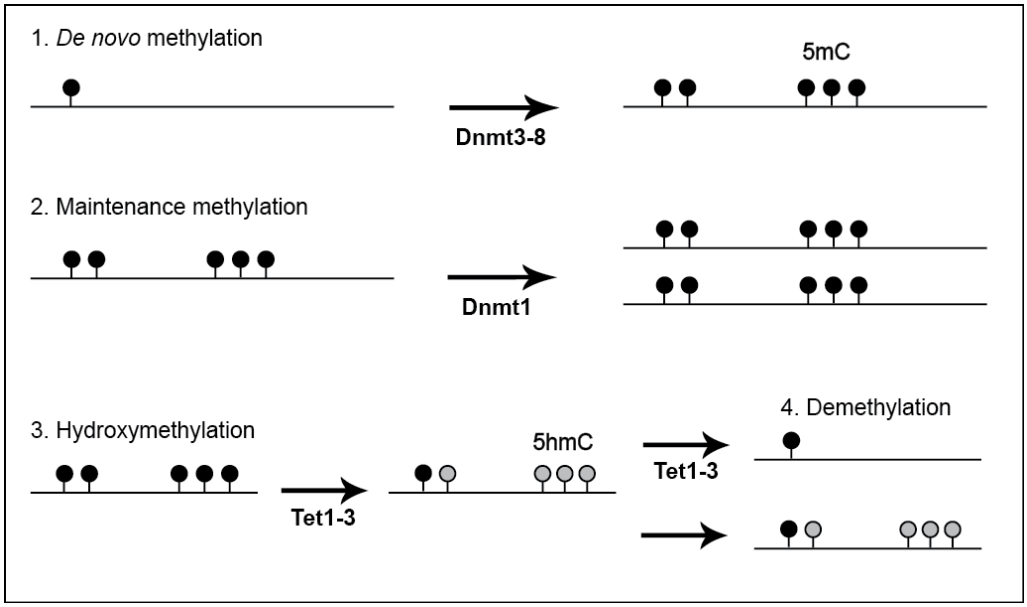

Figure 1.5: DNA methylation pathway. *De novo* DNA methyltransferases (Dnmt3-8) add new methyl groups onto cytosine (5mC) which is maintained through DNA replication by Maintenance methyltransferase (Dnmt1). Methylcytosine dioxygenases (Tet1-3) oxidize 5mC into 5-hydroxymethyl cytosine (5hmC) which can then be converted back into C (demethylated).

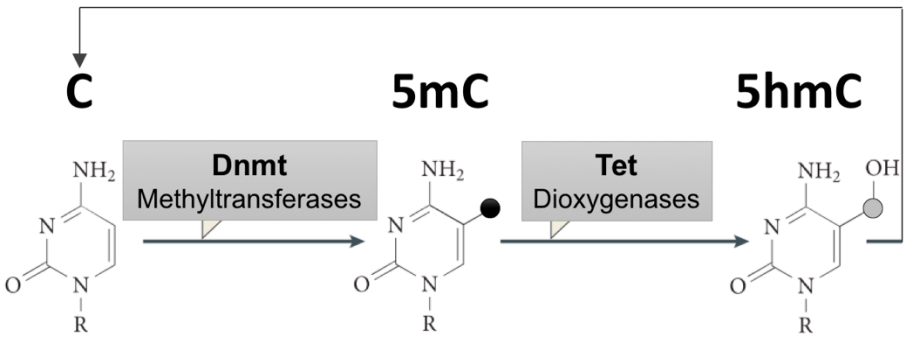

Figure 1.6: 5mC and 5hmC cytosine modifications are generated by Dnmt and Tet enzymes, respectively.

1.4 DNA METHYLATION AND HYDROXYMETHYLATION DURING DEVELOPMENT

Evidence from studies within the past decade points to DNA methylation as an important mechanism for the regulation of development in various contexts. Inactivation of *de novo* and/or maintenance methyltransferases causes phenotype associated with tissue-specific gene expression in various organisms. In humans, somatic mutations in *DNMT3A* are associated with acute leukemia and certain types of solid cancer (Kim et al., 2013a). In mouse, *Dnmt3a* null embryos develop normally but die soon after birth, *Dnmt3b* null embryos display neural tube defects, and *Dnmt3a/b* double mutants show early and severe phenotypes such as growth and morphogenesis arrest and lack of somites (Okano et al., 1999; Smith and Meissner, 2013; Wu et al., 2012). Embryonic stem (ES) cell lines deficient in all dnmt3-family genes exhibit global hypomethylation and premature differentiation into glial-like cells (Wu et al., 2012). Within the eye, mutations in *dnmt1*, and its recruitment factor, *uhrf1*, are required for normal development and maintenance of the zebrafish lens, and loss-of-function mutations in either results in reduced proliferation and loss of tissue integrity of the LE (Tittle et al., 2011). Tissue-specific combinatorial mutants for *Dnmt1*, *Dnmt3a* and *Dnmt3b* in mouse display global hypomethylation, reduction in terminally differentiated PRs, and disorganized synaptic formation (Singh et al., 2016).

Similar to *Dnmt*-family genes, mutations in *Tet* genes also cause a multitude of developmental defects. In mouse, *Tet1* knockout embryos show forebrain defects at late gastrulation and high mortality (Khoueiry et al., 2017); *Tet2* knockouts develop normally and are fertile, but *Tet1*$^{-/-}$*;Tet2*$^{-/-}$ double knockouts show a perinatal lethality associated with imprinting abnormalities (Dawlaty et al., 2013); *Tet1/2/3* triple knockout ES cells possess a massive loss of 5hmC, deregulated gene expression, and are less competent to differentiate (Dawlaty et al., 2014). Depletion of *tet3* by morpholino (MO) causes an

eyeless phenotype in Xenopus (Xu et al., 2012). In zebrafish, *tet2* MO knockdown results in mild erythropoiesis defects (Ge et al., 2014). However, somatic mutations in *tet2* do not cause any overt embryonic phenotype, although *tet2*$^{-/-}$ adults develop progressive age-related clonal myelodysplasia (Gjini et al., 2015). A recent study shows overlapping roles for *tet2* and *tet3* during hematopoietic stem cell differentiation (Li et al., 2015).

Within the nervous system, *Tet* expression and 5hmC enrichment is detected in the developing vertebrate brain and retina (Almeida et al., 2012b; Perera et al., 2015), and this coincides with the increase in 5hmC levels during neuronal differentiation at enhancers and also within gene bodies of neuronal genes (Hahn et al., 2013). Overexpression of *Tet3* in mouse olfactory neurons results in increased 5hmC levels and defects in axon targeting (Colquitt et al., 2013). Although both DNA methylation and hydroxymethylation have been studied extensively in the context of hematopoiesis, stem cell programming and disease, far less is known about their roles during development and in regulating cellular differentiation to create a complex structure like the eye.

1.5 Zebrafish as a Model for Eye Developmental Epigenetics

Zebrafish is a powerful vertebrate model system with unique attributes preferable for studying complex biological processes such as the epigenetic control of eye development, these include: amenability to genetic and biochemical manipulation, high fecundity, transparent and rapidly developing embryos, fully sequenced and well annotated genome, and extensive availability of recombinant constructs, mutants, and transgenic lines. Minor differences exist between zebrafish and human eye development. In zebrafish, the lens placode delaminates from the surface ectoderm as a solid mass of cells, rather than as a hollow vesicle as it does in mammals (Greiling and Clark, 2009;

Greiling et al., 2010). In the retina, the CMZ enables zebrafish retinae to continuously grow throughout life and regenerate in response to an injury, a phenomenon not observed in mammals (Raymond et al., 2006). Beyond these differences, all other aspects of lens and retina development appear to be identical between zebrafish and mammals, and the overall architecture of zebrafish and mammalian retinae and lenses are virtually identical (Bibliowicz et al., 2011; Gestri et al., 2012). This suggests that well conserved genetic and epigenetic mechanisms regulate vertebrate eye development and growth. Given its amenability to forward and reverse genetics (Auer and Del Bene, 2014; Bedell et al., 2012a; Cade et al., 2012; Donato et al., 2016; Hwang et al., 2013; Lee et al., 2012), zebrafish is an attractive model system in which to study the epigenetic regulation of eye development.

The goals of my thesis research were to investigate the roles of DNA methylation and hydroxymethylation during eye development, using zebrafish as a model organism. Utilizing a combination of classical embryology techniques, targeted genome editing, and genome-wide profiling approaches, I performed functional tests for genes within the *dnmt3* family (Chapter 2) and *tet* family (Chapter 3) during eye development, and generated the first combined methylome and hydroxymethylome maps of RPCs during retinal neurogenesis (Chapter 4).

Chapter 2: Gene expression and Functional Analysis of the *de novo* DNA Methyltransferases (*dnmt3-dnmt8*) during eye development

Portions of this Chapter are modified from the following article, with permission from the authors: Seritrakul, P., Gross, J.M., 2014. Expression of the De Novo DNA Methyltransferases (dnmt3 – dnmt8) During Zebrafish Lens Development. Developmental Dynamics. 243:350–356.

Seritrakul, P. and Gross, J.M. conceived the experiments and interpreted the data. Seritrakul, P. performed the experiments, collected and analyzed the data.

2.1 INTRODUCTION

De novo DNA methylation, mediated by Dnmt3a and Dnmt3b in mammals, is a process in which unmethylated CpG sites become methylated, and has been shown to play critical roles during development of plants and animals, particularly in tissue differentiation when cells undergo epigenetic reprogramming to adopt new identities (Feng et al., 2010; Hu et al., 2012; Illingworth et al., 2008; Merbs et al., 2012). Defects in, or loss of, *de novo* methyltransferase function cause a multitude of phenotypes associated with aberrant tissue-specific gene expression, and in a variety of animal model systems (e.g., (Gao et al., 2011; Okano et al., 1999; Wu et al., 2012)). Additionally, in humans, somatic mutations in DNMT3A are associated with acute leukemia and certain types of solid cancer (Kim et al., 2013b).

Previous work in our laboratory has shown that the functions of maintenance methyltransferase, dnmt1, and its recruitment factor, uhrf1, are required for normal development and maintenance of the zebrafish lens (Tittle et al., 2011). Both genes are expressed in an overlapping domain in the lens epithelium that correlates with regions of proliferation, and loss-of-function mutations in either results in altered LE gene expression, reduced proliferation, and loss of tissue integrity in the lens epithelium. While these data support a role for the maintenance methyltransferase dnmt1, during lens

development, little is known regarding the roles of *de novo* DNA methylation during vertebrate eye development.

Zebrafish possess six "dnmt3 family" *de novo* methyltransferase genes, *dnmt3 – dnmt8* (reviewed in (Goll and Halpern, 2011)). This group includes both orthologs of mammalian *dnmt3A* and *dnmt3B* as well as fish-specific genes with no obvious mammalian orthologs (Campos et al., 2012; Shimoda et al., 2005). Despite our understanding of their evolutionary history and their likely functional importance in regulating de novo methylation during organ and tissue development, the tissue-specific expression of these dnmt3-family genes is largely unknown. This is particularly true in the vertebrate eye, where we have a poor understanding of how DNA methylation regulates retina and lens development.

In mammals, loss-of-function alleles have already been generated: *Dnmt3a* null mouse embryos develop normally but die soon after birth; *Dnmt3b* null mutants display neural tube defects and die before birth; and *Dnmt3a/b* double mutants show a more severe early phenotypes than single mutants (Okano et al., 1999; Smith and Meissner, 2013; Wu et al., 2012). However, these early defects occur prior to eye development, thus preventing analysis of ocular phenotypes. Mouse embryonic stem cell lines deficient in *de novo* methyltransferases exhibit global hypomethylation, increased proliferation and precocious differentiation into glial-like cells (Wu et al., 2012). However, these phenotypes may not reflect dnmt3-specific phenotypes in developing vertebrate eyes in vivo. Functional studies in non-mammalian model organisms such as frog and zebrafish have so far relied heavily on morpholino-mediated knockdown of single *dnmt* genes, a relatively ineffective approach for interrogating families of proteins with potential for functional redundancy (Rai et al., 2010; Shimoda et al., 2005).

To facilitate functional studies of *de novo* methylation in the lens and retina, we have catalogued the distribution of dnmt3- family genes in the developing zebrafish eye over time and generated single and double knockout for dnmt3-family genes. To investigate the roles of *de novo* DNA methylation during zebrafish eye development, we utilized gene targeting strategies to create site-specific loss-of-function mutations in dnmt3-family genes that are expressed during the development of zebrafish lens and retina. Our data demonstrate that dnmt3-family genes are expressed in both unique and overlapping patterns, and in several ocular structures, supporting a model in which they likely play functional roles in regulating numerous aspects of normal eye development.

2.2 RESULTS

2.2.1 Phylogenetic Analysis of Zebrafish *De Novo* Methyltransferases

To examine the phylogenetic relationship among DNA methyltransferases in zebrafish and mammals, we constructed a phylogenetic tree using zebrafish, mouse, and human *de novo* and maintenance DNA methyltransferase proteins. Using dnmt2 RNA methyltransferase as an outgroup, *de novo* and maintenance methyltransferases cluster into two distinct clades, with all Dnmt1 orthologs in one and all Dnmt3 proteins in the other. Within the dnmt3 clade, dnmt3 and dnmt5 form a distinct branch with no mammalian ortholog, while dnmt6 and dnmt8 closely cluster with mouse Dnmt3A, consistent with previous studies (Campos et al., 2012; Goll and Halpern, 2011; Shimoda et al., 2005). The rest of the *de novo* methyltransferase clade includes zebrafish dnmt4, dnmt7, and mouse Dnmt3B. Within this group, dnmt4 branches with mouse Dnmt3B, while dnmt7 does not (Fig. 2.1). Therefore, zebrafish dnmt3, dnmt5, and dnmt7 appear to be fish-specific with no obvious mammalian orthologs.

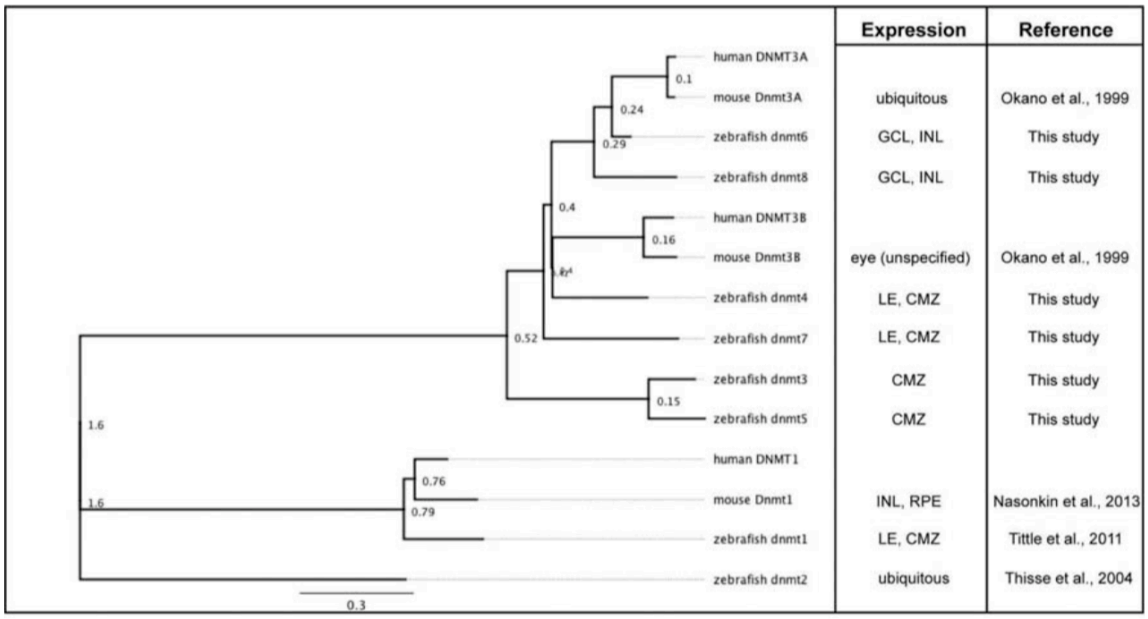

Figure 2.1: Phylogenetic relationship between DNA methyltransferases and their corresponding expression patterns. Tree was constructed using Geneious Tree Builder from an alignment of DNA methyltransferase amino acid sequences (zebrafish dnmt1, 3-8; mouse and human Dnmt1, 3A and 3B), with zebrafish RNA methyltransferase (dnmt2) as an outgroup.

2.2.2 Expression of Zebrafish De Novo Methyltransferases

All dnmt3-family genes show ubiquitous expression throughout the anterior region of the embryo at 24 hr postfertilization (hpf), after which, expression patterns begin to resolve into distinct, spatially restricted domains. Within the developing eye, expression domains can be categorized into three general groups, with each group constituted by two dnmt3-family genes (summarized in Fig 2.5).

2.2.2.1 CMZ expression: dnmt3 and dnmt5

At 24 hpf, *dnmt3* and *dnmt5* are expressed in the eye, as well as broadly throughout the embryonic head, and weakly in the posterior part of the embryo (Fig. 2.2A,F). Beginning at 36 hpf, and lasting through 72 hpf, transcripts of both genes are no longer detected throughout most of the embryo except for the ciliary marginal zone (CMZ) of the retina, where expression remains relatively high (Fig. 2.2B– D,G–I), and the optic tectum and pectoral fin. Transverse sections through the eyes highlight the CMZ expression of these two genes, and this is most pronounced at 48 hpf (Fig. 2.2K,L). At 4 days postfertilization (dpf), transcripts are still detected in the CMZ (Fig. 2.2E,J), but are no longer apparent in other parts of the eye or embryo. *dnmt3* and *dnmt5* transcripts are not detected within the lens.

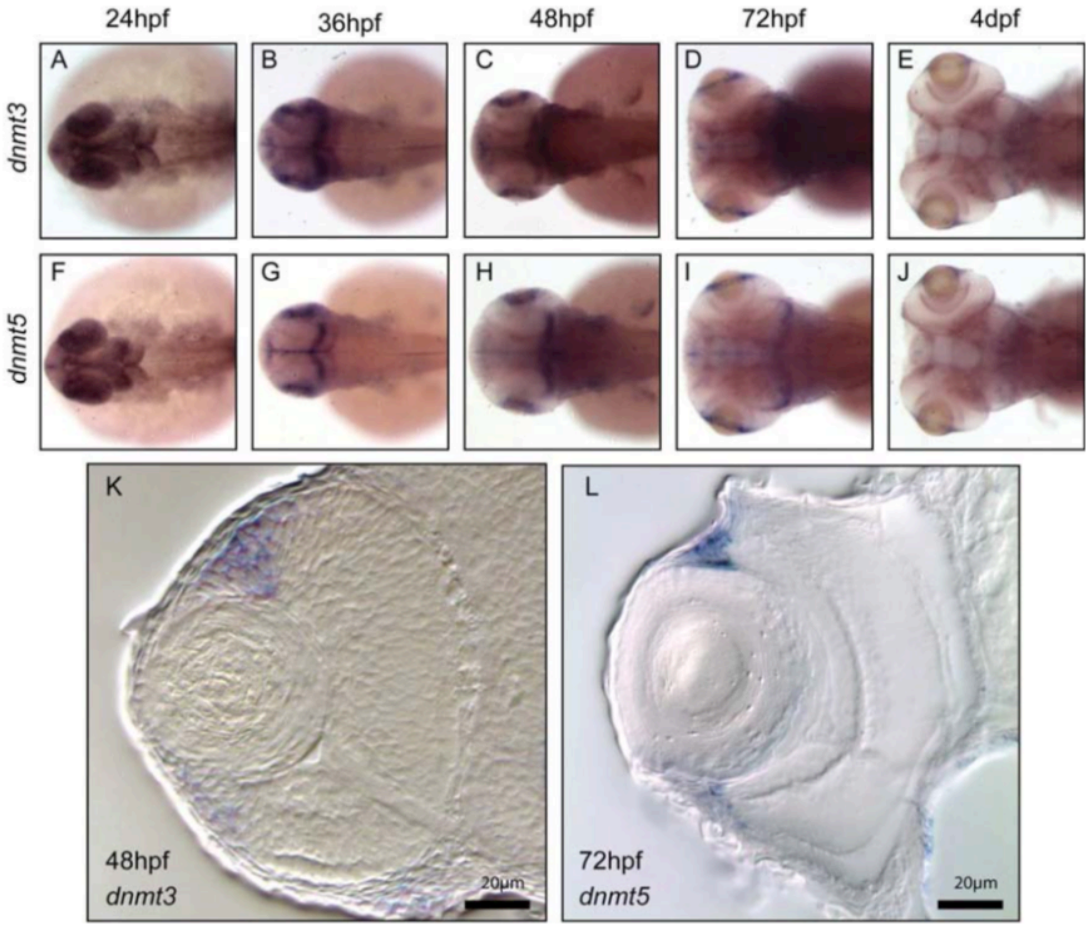

Figure 2.2: *dnmt3* and *dnmt5* expression is enriched in the ciliary marginal zone**.** A,F: In situ hybridization shows *dnmt3* and *dnmt5* expression throughout the anterior region of the embryo at 24 hr postfertilization (hpf). B–D,G–I: Expression becomes restricted to the CMZ, tectum, pectoral fins, and pharyngeal arches between 36 and 72 hpf. E,J: At 4 days postfertilization (dpf), ocular expression is only detected in the CMZ. Transverse sections, dorsal up (K,L).

2.2.2.2 CMZ and LE expression: dnmt4 and dnmt7

At 24 hpf, *dnmt4* and *dnmt7* expression resembles that of *dnmt3* and *dnmt5*, where transcripts are observed broadly throughout the anterior part of the embryo (Fig. 2.3A,F). In the eye, both genes are expressed within the CMZ, and throughout the lens epithelium (Fig. 2.3K,L). From 36 hpf to 4 dpf, *dnmt4* is also expressed in the optic tectum, pharyngeal arches, and posterior border of the developing pectoral fins (Fig. 2.3B–E). From 36 hpf on, outside of the eye, *dnmt7* expression appears to be overlapping but more diffuse than that of *dnmt4* (Fig. 2.3G–J).

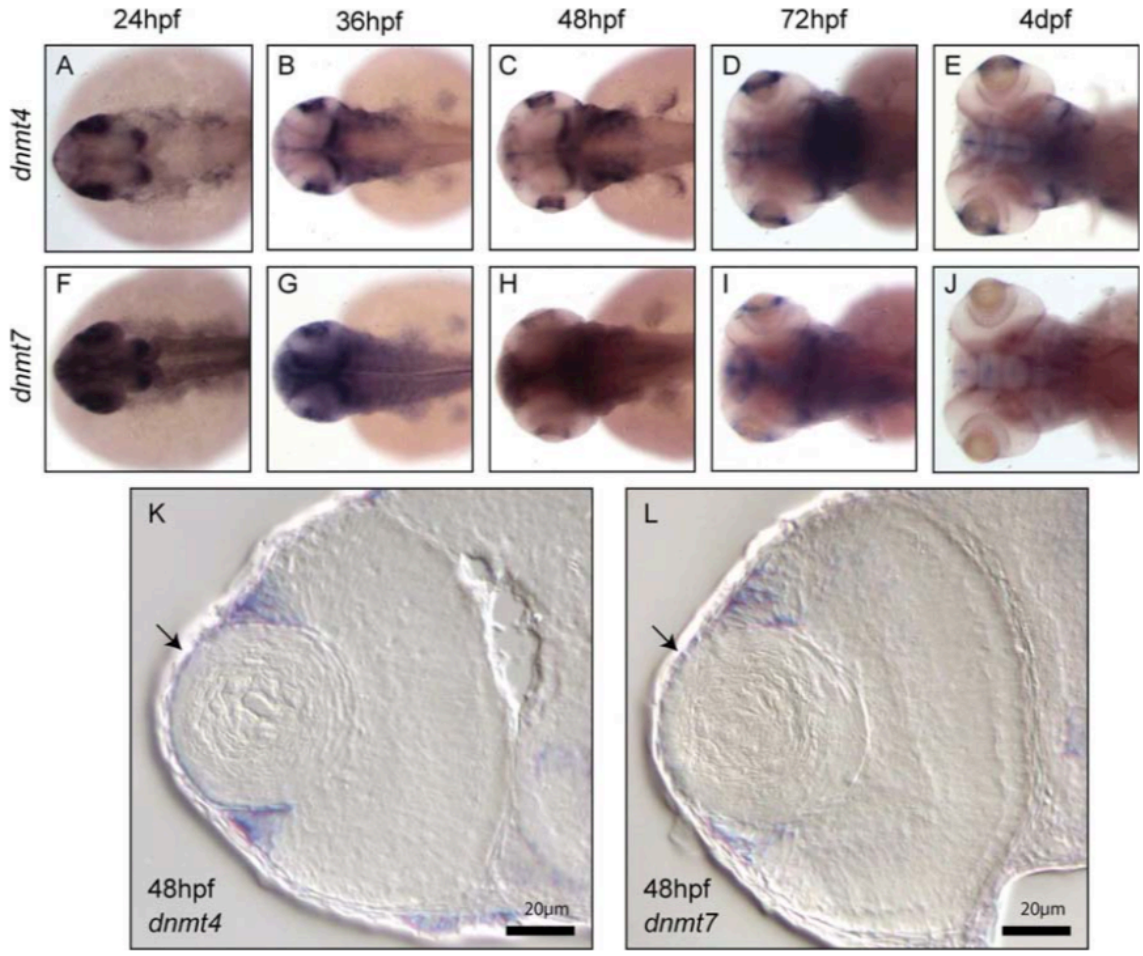

Figure 2.3: *dnmt4* and *dnmt7* expression in the lens epithelium and ciliary marginal zone. A–E,K: In situ hybridization shows broad *dnmt4* expression in the developing eye at 24 hr postfertilization (hpf) (A), which later becomes localized in the CMZ and lens epithelium (B–E, K, arrow points to LE). *dnmt4* expression is also observed in the tectum, developing pectoral fin, and pharyngeal arches (A–E). F: *dnmt7* is expressed ubiquitously at 24 hpf. G–J,L: At 36 hpf, expression remains diffuse in the anterior part of the embryo, and surrounding of the eye, while expression within the eye is only observed in the CMZ and LE (L, arrow points to LE). Transverse sections, dorsal up (K,L).

2.2.2.3 Inner Retina (GCL and INL) expression: dnmt6 and dnmt8

dnmt6 and *dnmt8* share a similar and overlapping expression pattern throughout embryogenesis, with both genes expressed ubiquitously at 24 hpf (Fig. 2.4A,F). Expression remains broad throughout the anterior of the embryo outside of the eye up to 4dpf (Fig. 2.4B– E,G–J). Within the eye, the expression appears to be dynamic. At 48 hpf, both genes are weakly expressed in the LE, CMZ, and developing ganglion cell layer (GCL) and inner nuclear layer (INL; Fig. 2.4C,H,K,L). However, at 72 hpf, expression is detected in only the GCL and the INL, and is no longer detected in the lens or CMZ (Fig. 2.4D,I,M,N).

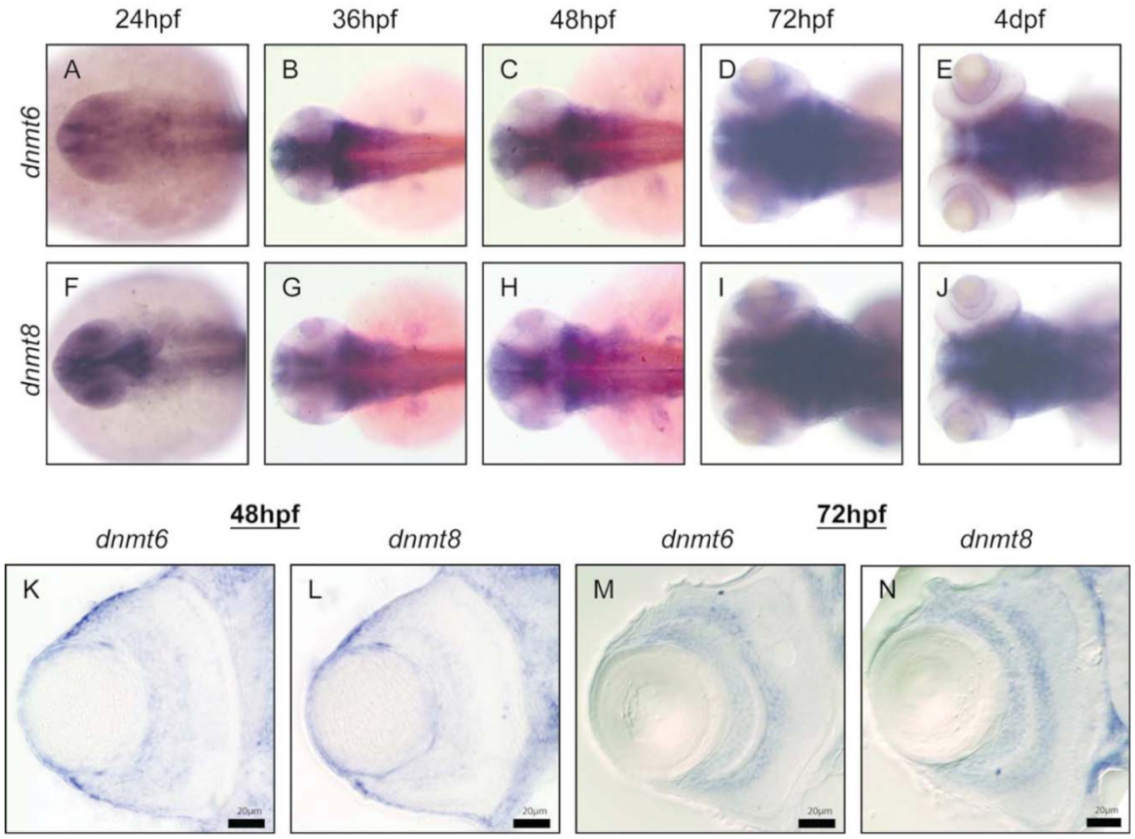

Figure 2.4: *dnmt6* and *dnmt8* expression in lens epithelial (LE), ciliary marginal zone (CMZ), ganglion cell layer (GCL), and inner nuclear layer (INL). In situ hybridization shows similar expression patterns between *dnmt6* and *dnmt8*. A,B,F,G: At 24–36 hr postfertilization (hpf), both genes are expressed ubiquitously throughout the anterior part of embryo. C,H,K,L: At 48 hpf, expression is detected specifically in LE and CMZ, and weakly in the developing GCL and INL. D,I,M,N: At 72 hpf, expression is only detected within the GCL and INL. E,J: At 4dpf, expression is faint in the GCL. K-N: Transverse sections, dorsal up.

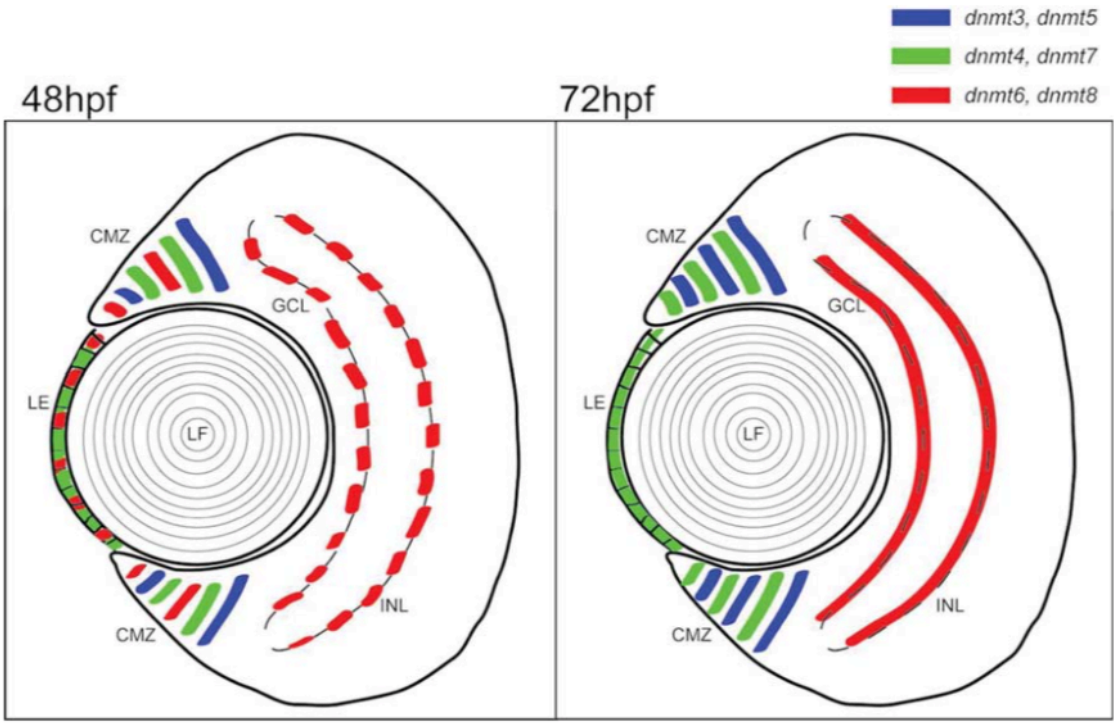

Figure 2.5: Summary of dnmt3-family gene expression in the developing zebrafish eye. Cartoon depicting a transverse section through an eye of a 48–72 hr postfertilization (hpf) embryo. Expression domains are marked in blue for dnmt3 and dnmt5, green for dnmt4 and dnmt7, and red for dnmt6 and dnmt8. LE, lens epithelial cells; LF, lens fibers; CMZ, ciliary marginal zone; GCL, ganglion cell layer; INL, inner nuclear layer.

2.2.3 Functional analysis of Zebrafish *De Novo* Methyltransferases

2.2.3.1 Characterization of dnmt4 and dnmt7 null mutants

To investigate roles of *de novo* methyltransferases in LE and CMZ, we generated site-specific heritable mutations in *dnmt4* and *dnmt7* using transcription activator-like effector nucleases (TALENs), a high-efficiency genome editing tool widely utilized in many model organisms (Carlson et al., 2012; Huang et al., 2011; Miller et al., 2011; Sakuma et al., 2013). TALENs were targeted to within a short distance from the start codon of *dnmt4* and *dnmt7*, mutations were validated by sequencing and are predicted to cause frame shift and premature stop codon, resulting in truncated, non-functional mutant proteins. Moreover, to eliminate the possibility of a second, alternate start codon re-initiating translation of the mutant protein, we also generated an additional *dnmt4* whole-gene knockout line utilizing two CRISPR guide RNAs (Auer et al., 2014; Irion et al., 2014). This mutant carries a large deletion between exon 3 and exon 22, knocking out the majority of the gene coding sequence, rendering the gene non-functional (Fig 2.6).

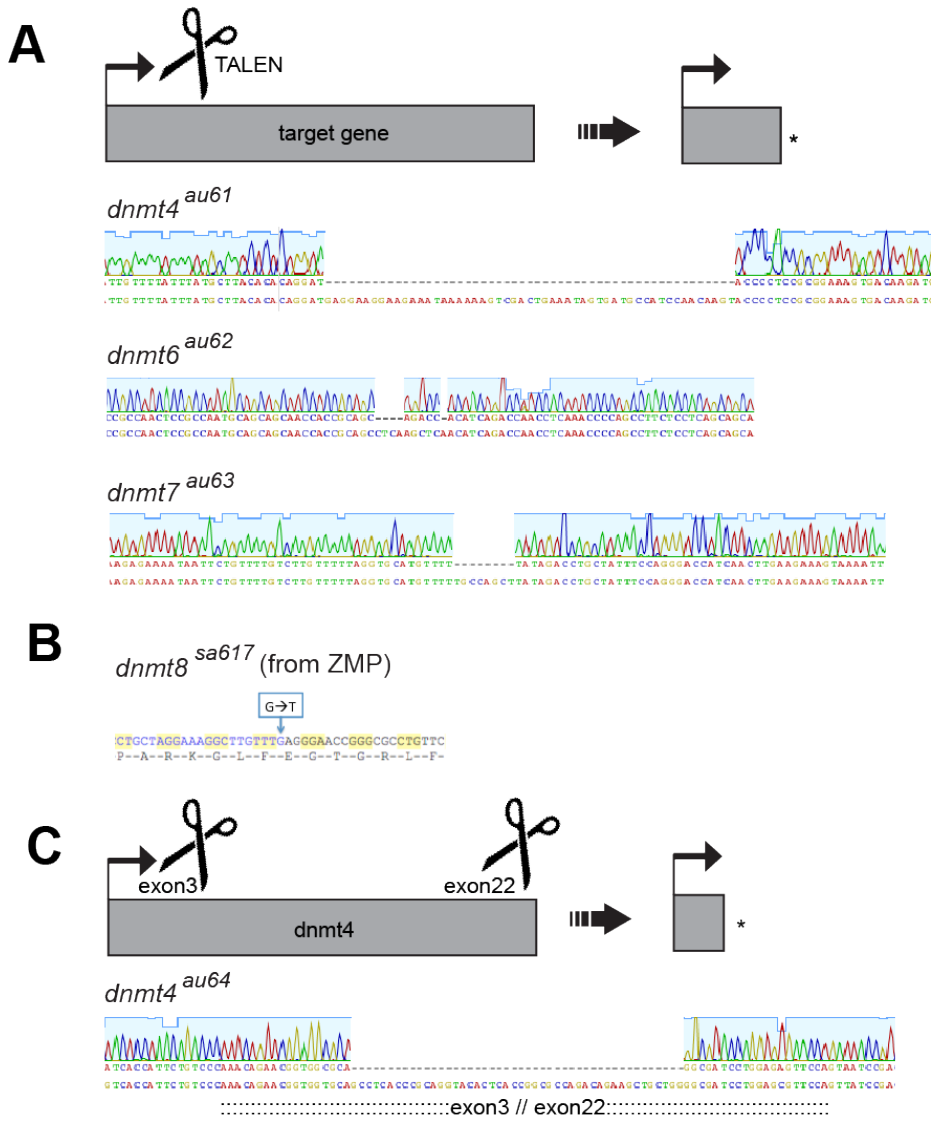

Figure 2.6: Generation of dnmt3-family mutant zebrafish lines. Single mutant lines were generated using TALENs, targeting sequences shortly downstream of the start codon. Mutant sequences for *dnmt4*, *dnmt6*, and *dnmt7* show out-of-frame deletion and are predicted to be null alleles (A). dnmt8 mutant lines obtained from Zebrafish Mutation Project shows a G→T point mutation that introduces an early stop codon (B). Whole-gene deletion of *dnmt4* shows large deletion from exon 3-22 (C). Chromatograms were obtained by amplicon Sanger sequencing.

Mutant embryos were individually genotyped by restriction fragment length polymorphism (RFLP) and recovered at the expected Mendelian ratio (25%), but the overall phenotype is indistinguishable from wildtype and heterozygous siblings. Further analysis of cryosectioned and immunostained eyes revealed normal development up to at least 7dpf for both *dnmt4* and *dnmt7* (Fig 2.7). Because mutant embryos carrying mutations in either *dnmt4* or *dnmt7* alone appear to develop normally and survive past early stages of development, and *dnmt4* and *dnmt7* are both expressed nearly identically in the CMZ and LE, we reasoned that these two proteins have redundant functions and thus compensate for each other in both mutant lines. To address this, we generated double heterozygous zebrafish, which were incrossed to generate *dnmt4/7* double mutant embryos at expected frequency of 6.25% (1 out of every 16 embryos). Double mutant also showed normal eye and general development and are viable to adulthood. Thus, *dnmt4* and *dnmt7* are likely not essential for normal development, and other dnmt3-family genes may compensate for their function in LE and CMZ.

2.2.3.2 Characterization of dnmt6 and dnmt8 null mutants

To investigate roles of *de novo* methyltransferases in the inner retina (GCL and INL), we generated a zebrafish mutant line carrying loss-of-function alleles in the *dnmt6* gene using TALENs as described. For *dnmt8*, a loss-of-function allele (*dnmt8*^{sa617}) was independently isolated by TILLING screening approach in the Zebrafish Mutation Project by the Sanger Institute but was not yet characterized (Kettleborough et al., 2013). We obtained this line and characterize the mutants for eye-specific phenotypes.

Similar to the scenario in *dnmt4* and *dnmt7*, analysis of cryosectioned and immunostained of either *dnmt6* or *dnmt8* single mutant embryos or dnmt6/8 double mutant embryos showed no eye-specific defects in overall lens or retinal morphology or cell type differentiation. All homozygous mutant larvae are viable, and adult mutants are fertile. Thus, *dnmt6* and *dnmt8* mutations do not appear to affect normal development or growth, again highlighting genetic redundancy between multiple genes within *de novo* methyltransferase family.

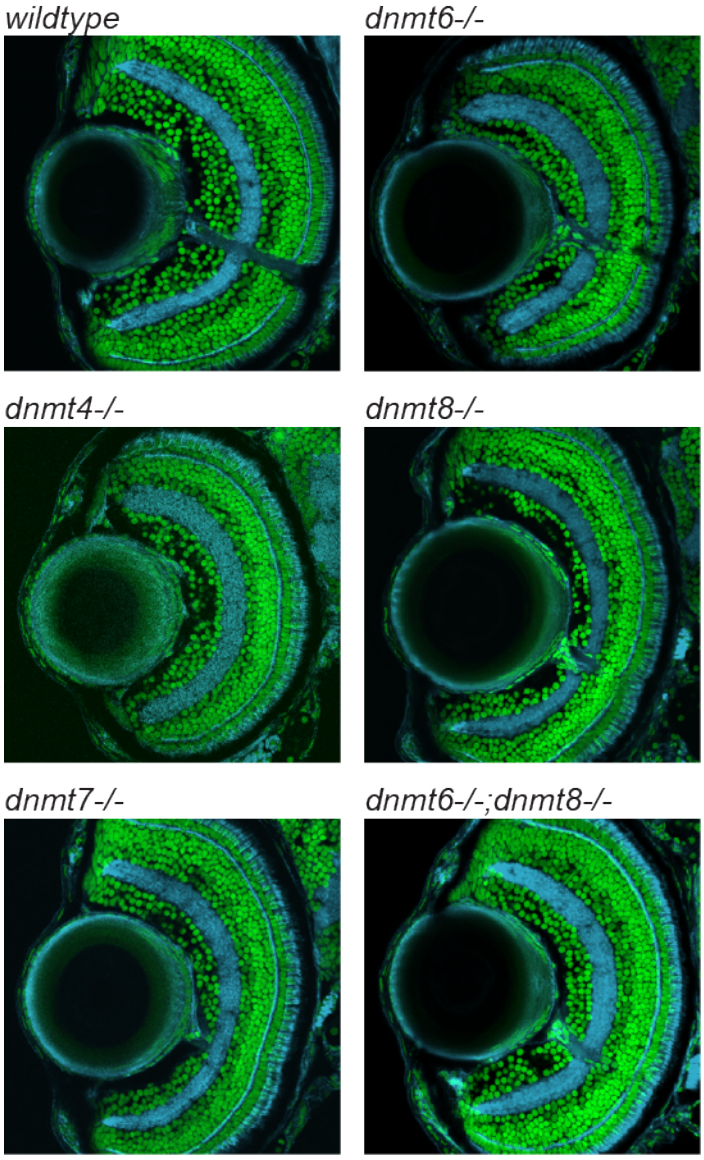

Figure 2.7: All dnmt3-family single and double mutants show normal eye development. 4dpf mutants were individually genotyped, sectioned, and stained with Sytox green for nuclei (green) and Phalloidin for F-actin (blue). All single and double homozygous mutants show normal eye morphology without overall developmental defect (n>3 embryos each).

2.3 DISCUSSION

Our *in situ* hybridization analysis reveals dynamic and spatiotemporally restricted patterns of *de novo* DNA methyltransferase gene expression during zebrafish eye development. These distributions suggest distinct, but possibly redundant, functions for dnmt3 family proteins during lens and retina development in vertebrates. Additionally, we generated and characterized mutants for dnmt-3 family genes, and double mutants for genes with overlapping expression domain. None of the single or double mutants show any developmental defect, indicating that these genes are highly functionally redundant, and can compensate for each other, even outside of their normal expression domains. This necessitate the generation of triple and/or quadruple mutants where all *de novo* methyltransferase function is abolished.

2.3.1 *De novo* DNA methyltransferases and Lens Development

The LE-specific expression of dnmt4/dnmt7 (from 36 hpf onward) and dnmt6/dnmt8 (at 48 hpf) suggest possible functions for the proteins they encode in 1) maintenance of the continuously proliferating LE cell population and/or 2) during the reprogramming of LE cells to become LF cells in the transition zone, a process that involves genome-wide changes in expression of numerous genes including activation of LF- specific genes and silencing of LE- specific genes (e.g., (Cvekl and Duncan, 2007; Hayes et al., 2012; Wride, 2011). For example, delta-crystallin genes (CRYD1/2) in the embryonic chicken lens undergoes changes in methylation status and transcription levels as LE cells differentiate into LFs (Sullivan and Grainger, 1986; Sullivan et al., 1989), suggesting that LE and LF cells are epigenetically distinct cell populations. Interestingly, these *de novo* methyltransferase transcripts are distributed throughout the entire LE, a domain that is broader than that of the maintenance methyltransferase (dnmt1), which is expressed solely in the proliferative regions of the LE (Tittle et al., 2011). Beyond the

regulation of gene expression during the LE to LF transition, *de novo* DNA methylation machinery could cooperate with chromatin-remodeling enzymes in the LE cells to mediate chromatin compaction in preparation for DNA degradation that occurs during LF terminal differentiation (Appleby and Modak, 1977; Bassnett et al., 2009; He et al., 2010). Lack of proper DNA degradation causes nuclear cataract in mouse (Nishimoto et al., 2003), and hypermethylation of a LF-specific aA-crystallin (CRYAA) contributes to age- related cataract in human (Zhou et al., 2012). These examples also suggest that defects in *de novo* DNA methylation may play a role in the progression of ocular diseases.

2.3.2 *De novo* DNA methyltransferases and Retinal Development

All six members of the dnmt3 gene family are expressed within the retina in two distinct domains: the CMZ and the GCL-INL layers. Because the CMZ is an area of active cell proliferation and a known niche for retinal stem cells (Raymond et al., 2006), it is possible that these de novo DNA methyltransferases function in maintaining the proliferative properties of CMZ cells and/or in the reprogramming of retinal precursors destined to exit the CMZ and differentiate into retinal neurons. Consistent with this possibility, a recent study in Xenopus demonstrated that the transcripts for all core components of the polycomb repressive complex 2 (PRC2), a chromatin remodeling complex that catalyzes histone H3 lysine 27 tri-methylation (H3K27me3), are also expressed within the CMZ (Aldiri et al., 2013). These data suggest that the CMZ is an area of highly active epigenetic modifications, and that *de novo* DNA methylation may cooperate with histone-modifying enzymes to facilitate stem cell maintenance or reprogramming during differentiation.

The distribution of *dnmt6* and *dnmt8* in the GCL and INL of the retina, as well as the absence of dnmt3 family gene expression in other regions of the differentiated retina suggest that GCL and INL cells may undergo additional epigenetic modification before they terminally differentiate. These two layers of the retina are also where 5-hydroxymethylcytosine (5hmC), a mark of active cytosine demethylation and gene activation (Pastor et al., 2013), is enriched (Almeida et al., 2012a). Recent findings suggest that Tet enzymes, which are responsible for 5mC to 5hmC conversion, are expressed in the eye and likely facilitate neurogenesis (Li et al., 2014; Santiago et al., 2014; Xu et al., 2012). Therefore, it is likely that Tet enzymes play roles in terminal differentiation of inner retinal neuron. Experiments in Chapter 3 directly test this hypothesis.

2.3.2 Functional Redundancy in zebrafish *De novo* DNA methyltransferases

Given our current understanding of the evolutionary history of genes in the dnmt3 family and the overlapping distributions reported here (Fig 2.5), it is very likely that the encoded proteins have redundant roles during zebrafish lens and retinal development, with genes that are most closely related phylogenetically sharing expression domains and possible functions. This overlap in expression also highlights potential difficulties in performing functional experiments due to the possibility of genetic redundancy within this group of genes. Therefore, it is not surprising to see that neither the single or double mutant for dnmt3-family genes show any developmental defects.

It is important to note that our double mutants were generated between pairs of dnmt3-family genes that are expressed in the same domains, in the hope that the lack of both gene in each domain will lead to a developmental phenotype. However, it is possible that, because these genes share nearly identical catalytic domains (Goll and

Halpern, 2011), they may compensate for each other even outside of their normal expression domain. For example, dnmt6/8 could compensate for dnmt4/7 in the LE, even though *dnmt6/8* transcripts are not normally detected there. Additionally, dnmt1 also possess a fully functional methyltransferase domain, and could also perform functional compensation. In mouse, recent finding demonstrate that tissue-specific triple knockout of *Dnmt1*, *Dnmt3A* and *Dnmt3B* lead to defects in the differentiation of photoreceptor and formation of outer plexiform layer (Singh et al., 2016). Moving forward, functional analyses will necessitate the generation of combinatorial loss-of-function animals where more than two dnmt3-family gene is simultaneously perturbed. Current work in our laboratory is aimed at simultaneously deleting or inactivating all members of the dnmt3-family, utilizing a combination of individual mutations and large chromosomal deletion encompassing dnmt3,4,5 cluster, but this is beyond the scope of my Dissertation

Chapter 3: Tet-mediated DNA hydroxymethylation regulates retinal neurogenesis by modulating cell extrinsic signaling pathways

Portions of this Chapter are modified from the following article, with permission from the authors: Seritrakul, P., Gross, J.M., 2017. Tet-mediated DNA hydroxymethylation regulates retinal neurogenesis by modulating cell-extrinsic signaling pathways. PLOS Genetics. 13(9): e1006987.

Seritrakul, P. and Gross, J.M. conceived the experiments and interpreted the data. Seritrakul, P. performed the experiments, collected and analyzed the data.

3.1 INTRODUCTION

Recent studies have identified roles for Tet proteins and DNA hydroxymethylation during vertebrate development, stem cell maintenance and in diseases such as cancer. In mouse, *Tet1* and *Tet2* knockouts are viable and fertile, while *Tet1*$^{-/-}$*;Tet2*$^{-/-}$ double knockouts show a partially penetrant perinatal lethality associated with imprinting abnormalities (Dawlaty et al., 2013). A recently generated Tet1 knockout mouse showed forebrain defects at late gastrulation and high mortality (Khoueiry et al., 2017). *Tet1/2/3* triple knockout ES cells possess a massive loss of 5hmC, deregulated gene expression, and an impaired ability to differentiate (Dawlaty et al., 2014). In frog, depletion of Tet3 by a translation-blocking morpholino (MO) affects early neural development and causes an eyeless phenotype (Xu et al., 2012). *tet2* MO-based knockdown in zebrafish results in mild defects in erythropoiesis (Ge et al., 2014); however, mutations in *tet2* do not cause any overt embryonic phenotype, although *tet2*$^{-/-}$ adults develop progressive age-related clonal myelodysplasia (Gjini et al., 2015). More recently, overlapping roles for *tet2* and *tet3* during hematopoietic stem cell differentiation have been identified (Li et al., 2015). Likely related to their functions during hematopoiesis, Tet proteins are also associated with a number of hematological malignancies in humans [reviewed in (Ko et al., 2015)].

Within the nervous system, Tet expression and 5hmC enrichment is detected in the developing mouse brain (Almeida et al., 2012b; Hahn et al., 2013). 5hmC levels increase during neuronal differentiation, with enrichment at enhancers and also within gene bodies of neuronal genes (Hahn et al., 2013). This enrichment is interesting

because it is not associated with subsequent demethylation, in agreement with a study that shows biochemical stability of 5hmC marks within the genome (Bachman et al., 2014). Beyond loss-of-function experiments, *Tet3* overexpression in mouse olfactory neurons results in an increase in 5hmC levels, altered gene expression, and defects in axon targeting (Colquitt et al., 2013), and tet3 activity is upregulated in dorsal root ganglia neurons during axon regeneration (Loh et al., 2016).

More recent work has shown that Tet proteins regulate both intrinsic and extrinsic pathways during development. Intrinsically, Tet activity is required during hematopoiesis (Madzo et al., 2014) and B-cell development (Orlanski et al., 2016). Tets also modulate the activity of extrinsic signaling pathways in a variety of contexts. Tets suppress Wnt pathway activity during early mouse development to balance mesoderm and neuroectoderm fates (Li et al., 2016), while at later stages in the intestinal epithelium, Tet1 is required for Wnt pathway activation (Kim et al., 2016). During gastrulation, Tet activity regulates the Nodal pathway by suppressing the expression of Nodal inhibitors (Dai et al., 2016).

Although DNA methylation and hydroxymethylation have been studied extensively in the context of stem cell programming and disease, far less is known about their roles during development and in regulating organogenesis to create a complex structure like the retina. The vertebrate retina consists of seven main cell types that perform distinct functions in phototransduction and visual signal transmission [reviewed in (Bassett and Wallace, 2012)]. These cells are organized in the three retinal layers: the ganglion cell layer (GCL), outer nuclear layer (ONL), and inner nuclear layer (INL). The retina develops from a common pool of seemingly indistinguishable multipotent retinal progenitor cells (RPCs), which ultimately give rise to all retinal cells in a stereotyped time order [reviewed in (Centanin and Wittbrodt, 2014)].

While much has been learned regarding the specification events that direct RPCs to distinct retinal cell fates, we know little about the epigenetic regulation of these events, or the mechanisms underlying terminal differentiation and morphogenesis of retinal neurons and glia. Indeed, no studies have determined whether Tet function or DNA

hydroxymethylation are required during retinal development. To address this topic we functionally inactivated the *tet2* and *tet3* genes in zebrafish and identified defects in retinal development that resulted from deregulated gene expression in *tet2*$^{-/-}$;*tet3*$^{-/-}$ mutants. This is the first detailed analysis of Tet function during vertebrate retinal development, and our data support a model in which Tet-mediated regulation of 5hmC levels is critical for retinal neurogenesis.

3.2 RESULTS

3.2.1 *tet2* and *tet3* expression and functional inactivation

Several studies have reported the expression of Tet-family genes in developing embryos and in tissues, including the eye (e.g. (Almeida et al., 2012b; Xu et al., 2012)). Zebrafish possess three Tet-family genes (*tet1*, *tet2* and *tet3*), and these are orthologous to mouse and human Tet1-3 (Fig. 3.1A). At 24 hours post fertilization (hpf) all three *tet* genes are expressed broadly throughout the embryo, including the eye (Fig. 3.1B-D). At 48 and 72hpf, *tet1* expression is faint in the head and is expressed at comparatively lower levels in the eye than *tet2* and *tet3* (Fig. 3.1E,H). At 48 and 72hpf, *tet2* and *tet3* are strongly expressed in the anterior of the embryo and, in the eye, transcripts are enriched in the GCL and INL, and are more faintly detected in the ONL (Fig 3.1F,G,I,J).

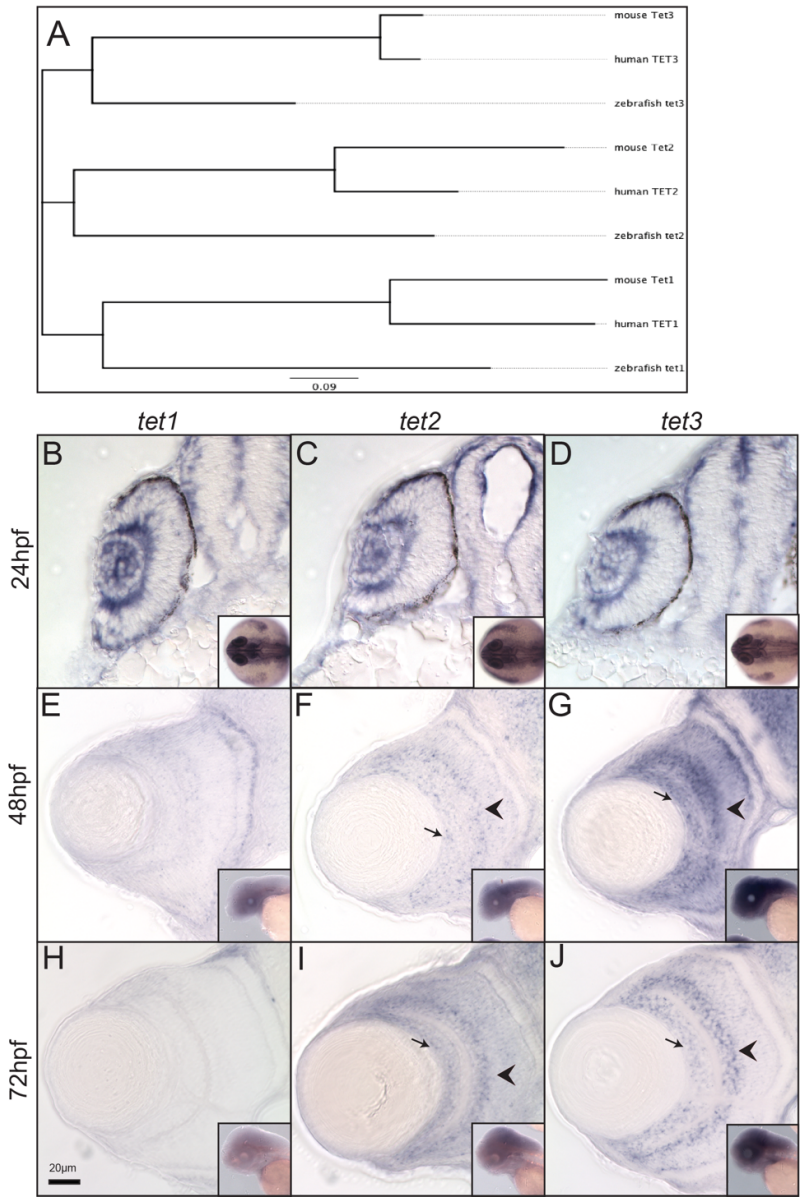

Figure 3.1: **Tet-family gene expression and phylogenetic analyses.** (A) An unrooted phylogenetic tree constructed from mouse, human and zebrafish Tet1, 2 and 3 proteins. (B-D) *tet1*, *tet2* and *tet3* are ubiquitously expressed at 24hpf. At 48 (E-G) and 72hpf (H-J) *tet2* and *tet3* are expressed in the inner nuclear layer (INL; arrowhead) and ganglion cell layer (GCL; arrows), and faintly in the outer nuclear layer (ONL). n>8 per gene per time point. Scale bar = 20µm.

Because *tet2* and *tet3* were the only Tet-family genes that showed prominent retinal expression past 24hpf, and a recent report demonstrated that *tet1* was dispensable for normal zebrafish development and for DNA hydroxymethylation (Li et al., 2015), we focused on *tet2* and *tet3* for functional perturbations. Mutant alleles were created by designing transcription activator-like effector nucleases (TALENs) targeting the first exon of *tet2* and the sixth exon of *tet3* (Bedell et al., 2012b). Injected mosaic founders were screened for germline transmission, and mutant alleles were detected and sorted by restriction fragment length polymorphisms (RFLP) and then sequenced (Fig. 3.2A,B; Fig 3.3A). Two alleles were identified and maintained: *tet2*^{au59} mutants possess a 10bp deletion which is predicted to cause a frameshift beginning at amino acid (aa) 407, inserting 37 incorrect amino acids and truncating the protein at aa 444 out of 1,716 [c.1,219_1,229delATAGATTTAA, p.Ile407Thrfs*37], and *tet3*^{au60} mutants possess a 22bp deletion mutation, which is predicted to cause a frameshift beginning at aa 1,212, inserting 92 incorrect aa and truncating the tet3 protein at aa 1,304 out of 2,052 [c.4,860_4,881delGAGATAAACTGTACAGAGAAGT, p.Gly1,212Alafs*92].

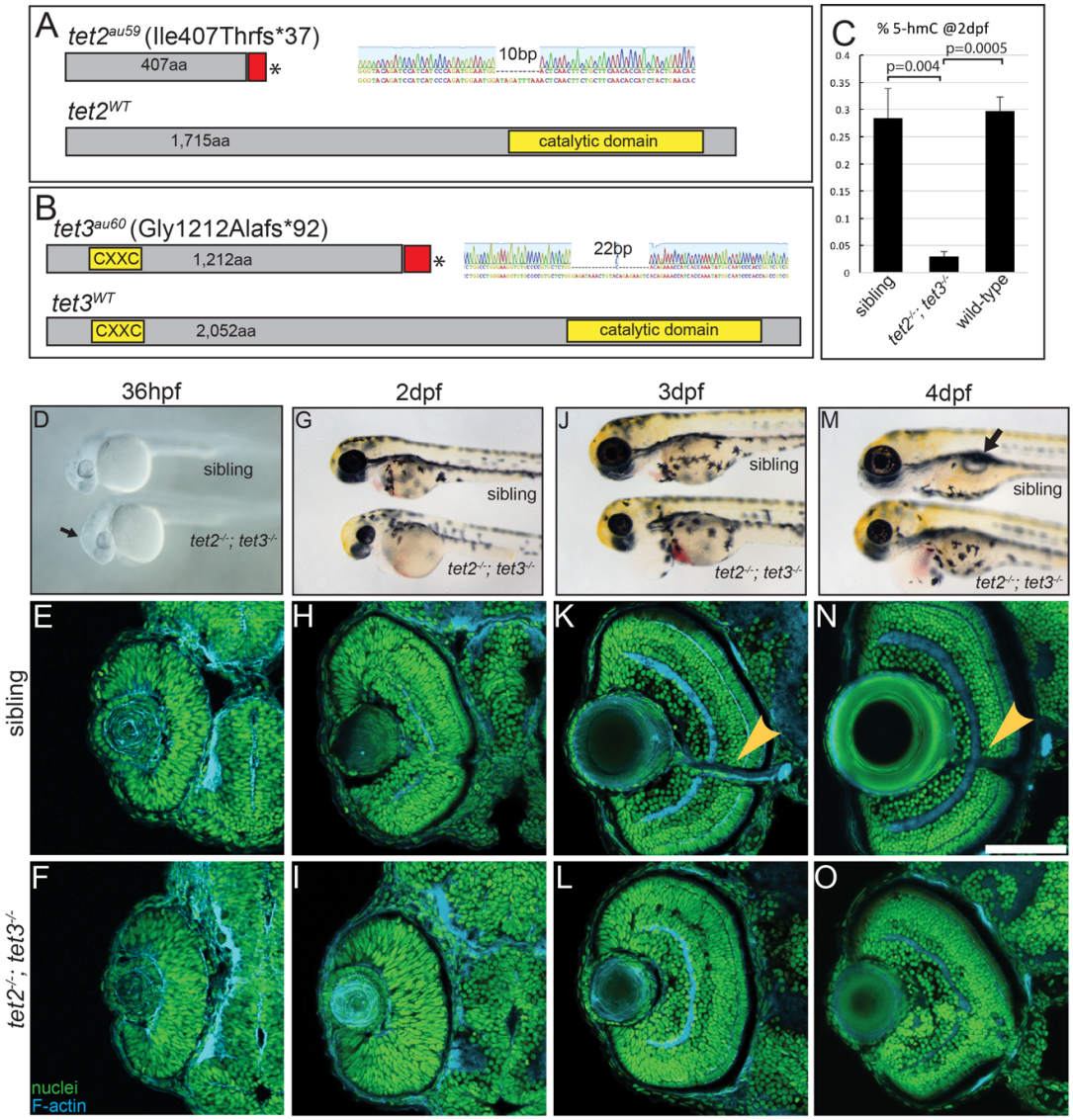

Figure 3.2: ***tet2*$^{-/-}$;*tet3*$^{-/-}$ mutants are deficient in 5mC 5hmC conversion and display abnormalities in retinal development.**

(Figure 3.2 - continued) (A) *tet2*^{au59} mutants possess a 10bp deletion resulting in a frameshift, insertion of 37 incorrect aa's (red), and a premature stop codon, truncating the protein at amino acid (aa) 444 of 1,716aa. (B) *tet3*^{au60} mutants possess a 22bp deletion, resulting in a frameshift, insertion of 92 incorrect aa's (red), and a premature stop codon, truncating the protein at aa 1,304 of 2,052. (C) Genomic DNA isolated from 2dpf *tet2*$^{-/-}$*;tet3*$^{-/-}$ mutants shows a >9 fold reduction in global 5hmC levels when compared to phenotypically wild-type siblings and genetically wild-type embryos in an ELISA (n=20 embryos, p=0.004 and p=0.0005, respectively; Error bars = ± 1 S.D.). (D) At 36hpf, *tet2*$^{-/-}$*;tet3*$^{-/-}$ mutants are identifiable based on a kinked head and slightly enlarged brain (arrow). (E,F) Retinae of 36hpf *tet2*$^{-/-}$*;tet3*$^{-/-}$ mutants are morphologically similar to wild-types. (G) At 2dpf, *tet2*$^{-/-}$*;tet3*$^{-/-}$ mutants are microphthalmic, possess cardiac edema and their heads are smaller than phenotypically wildtype siblings. (H-I) The retina is not laminated, with retinal cells appearing progenitor-like in morphology when compared to phenotypically wild-type siblings. (J) Cardiac edema becomes progressively enlarged at 3dpf in *tet2*$^{-/-}$*; tet3*$^{-/-}$ embryos. (K,L) The retina remains poorly laminated in *tet2*$^{-/-}$*;tet3*$^{-/-}$ mutants, and they lack a morphologically obvious optic nerve (arrowhead in sibling). (M) A 4dpf *tet2*$^{-/-}$*; tet3*$^{-/-}$ embryos do not possess an inflated swim bladder (arrow). (N,O) The retina remains poorly laminated and they lack a morphologically obvious optic nerve (arrowhead in sibling). DNA (green), F-actin (cyan). Dorsal is up and anterior to the left. Scale bar = 80μm.

tet2$^{-/-}$ and *tet3*$^{-/-}$ mutants develop normally with no visible phenotype and they are homozygous viable (Fig. 3.3B). This is an unsurprising result given their close phylogenetic relationship and overlapping expression domains, and it is consistent with recent reports (Gjini et al., 2015; Li et al., 2015). Thus, *tet2*$^{-/-}$*;tet3*$^{-/-}$ mutants were generated by in-crossing double heterozygous adults (*tet2*$^{+/-}$*;tet3*$^{+/-}$). *tet2*$^{-/-}$*;tet3*$^{-/-}$ mutants were recovered at an expected Mendelian ratio (6.62%; n=103 mutants/1,554 embryos). At 36hpf, *tet2*$^{-/-}$*;tet3*$^{-/-}$ mutants displayed a distinct morphological phenotype where the anterior portion of the brain was enlarged and kinked when compared to wild-type siblings (Fig. 3.2D). At 2 days post-fertilization (dpf), *tet2*$^{-/-}$*;tet3*$^{-/-}$ mutants were microphthalmic, mildly hypopigmented, and displayed deformed craniofacial features, cardiac edema, and blood pooling, phenotypes that perdure through 4dpf (Fig. 3.2G,J,M).

tet2^{au59} and *tet3*^{au60} mutations are predicted to truncate the proteins upstream of the C-terminal catalytic domain (Pastor et al., 2013) and therefore encode null or severe loss of function alleles. At 2dpf and 5dpf, *tet2* transcripts were detectable in phenotypically wild-type siblings, but not in *tet2*$^{-/-}$*; tet3*$^{-/-}$ mutants, indicating that *tet2* transcripts are degraded through nonsense-mediated decay (NMD) (Fig. 3.4A). *tet3* transcripts were still present in both sibling and *tet2*$^{-/-}$*;tet3*$^{-/-}$ mutants (Fig. 3.4B). To assess tet3 protein, Western blots were performed on 3dpf phenotypically wild-type siblings and *tet2*$^{-/-}$*;tet3*$^{-/-}$ mutants (Fig. 3.4C,D). tet3 protein was undetectable in *tet2*$^{-/-}$*;tet3*$^{-/-}$ mutants. These data indicate that *tet2*^{au59} and *tet3*^{au60} are null alleles. To experimentally validate the loss of catalytic function in *tet2*$^{-/-}$*;tet3*$^{-/-}$ mutants, we utilized a sensitive enzyme-linked immunosorbent assay (ELISA) to quantify whole embryo 5hmC levels. At 2dpf, genetically wild-type and phenotypically wild-type siblings possessed genome-wide 5hmC levels of 0.30% and 0.28%, respectively, consistent with published levels in various isolated mouse tissues (Khoueiry et al., 2017; Szwagierczak et al., 2010). By comparison, *tet2*$^{-/-}$*;tet3*$^{-/-}$ mutants possessed a significant reduction in 5hmC levels (0.03%, $p<0.005$) indicating that *tet2*^{au59} and *tet3*^{au60} mutants lack almost all 5hmC in their genome (Fig. 3.2C).

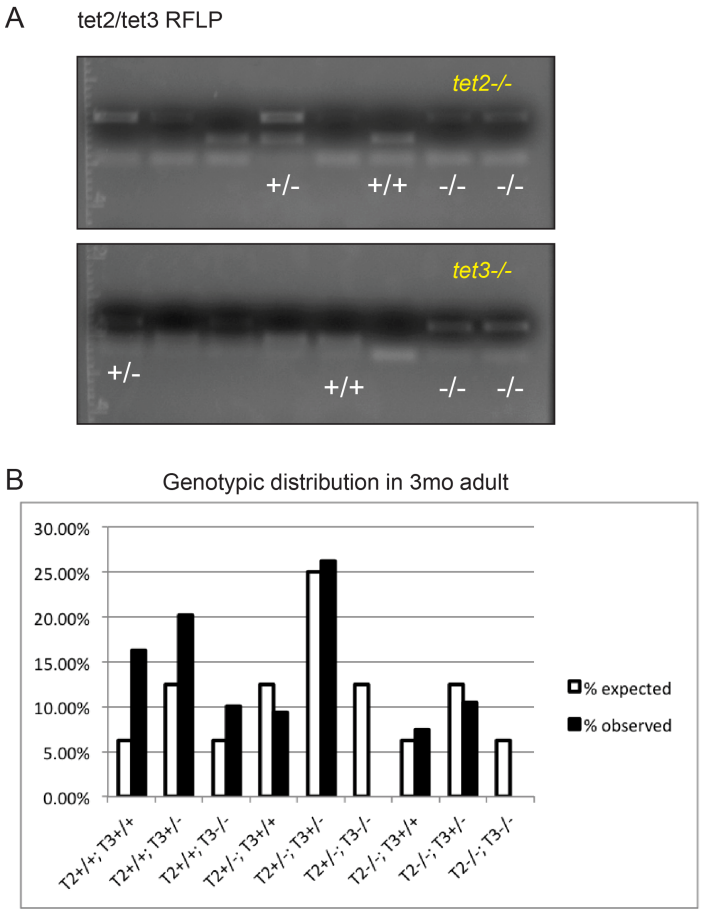

Figure 3.3: ***tet2* and *tet3* RFLP genotyping and genotypic distribution in adults. (A)** Mutations were detected by restriction fragment length polymorphism (RFLP). Mutant alleles lack the recognition site for DraI (for *tet2*) and RsaI (for *tet3*), and are therefore undigested. (B) At 3-months, 96 fish were individually genotyped by RFLP. The genotypic distribution follows a Mendelian distribution for a dihybrid cross, except for the absence of $tet2^{-/-}$;$tet3^{-/-}$ mutations, which are embryonic lethal and $tet2^{+/-}$;$tet3^{-/-}$ which are juvenile lethal.

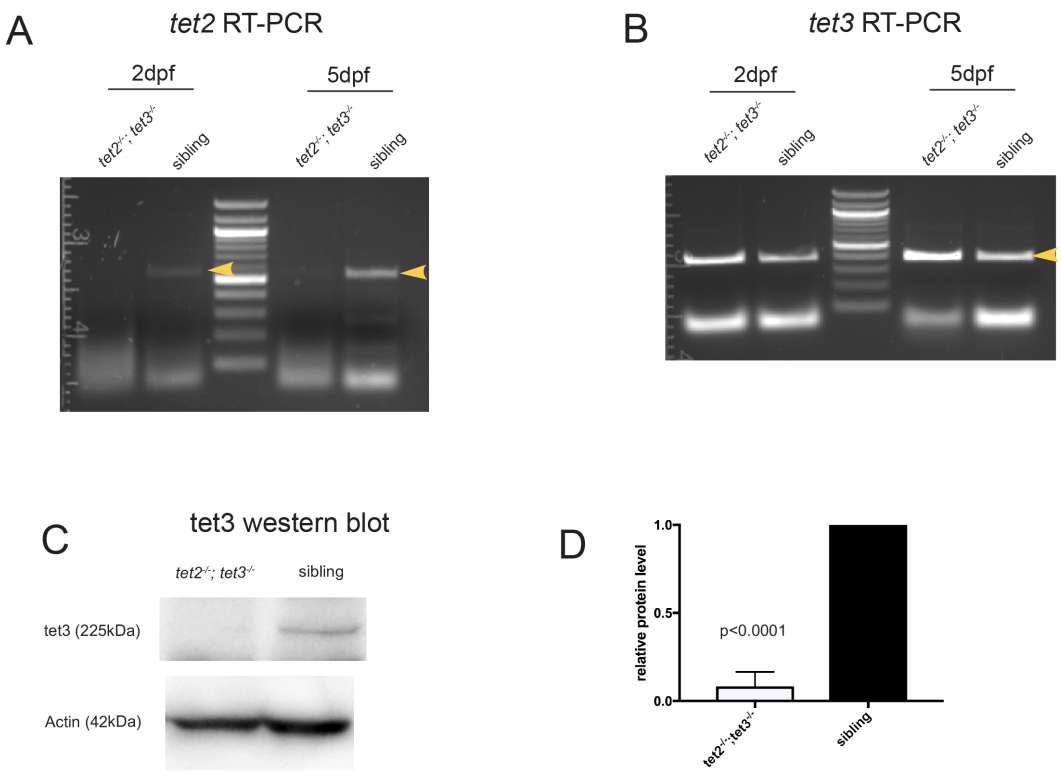

Figure 3.4: ***tet2* transcript and tet3 protein are undetectable in *tet2*$^{-/-}$;*tet3*$^{-/-}$ mutants.** (A) At 2dpf and 5dpf, *tet2* transcripts are present in sibling but undetectable by RT-PCR in *tet2*$^{-/-}$;*tet3*$^{-/-}$ mutants indicating degradation, presumably via nonsense-mediated decay. (B) *tet3* transcripts are present in both sibling and *tet2*$^{-/-}$;*tet3*$^{-/-}$ at both time points. N=20 embryos per condition, and experiments done in biological triplicates. RT-PCRs for *tet2* and *tet3* were done in parallel from the same cDNA pools. (C,D) At 3dpf, tet3 protein (225 kDa) is absent from *tet2*$^{-/-}$;*tet3*$^{-/-}$ mutants. N=40 embryos per condition, and experiments done in biological triplicates. P<0.0001, unpaired t-test.

To analyze retinal development in *tet2*$^{-/-}$;*tet3*$^{-/-}$ mutants, we performed histology from 36hpf to 4dpf and assessed overall retinal structure. At 36hpf, no obvious differences between retinal morphology in *tet2*$^{-/-}$;*tet3*$^{-/-}$ mutants and wild-type siblings were evident (Fig. 3.2E,F). At 2dpf, *tet2*$^{-/-}$;*tet3*$^{-/-}$ mutant retinae appeared progenitor-like in morphology, as little evidence of lamination or neuronal differentiation was evident (Fig. 3.2H,I). At 3dpf, *tet2*$^{-/-}$;*tet3*$^{-/-}$ mutant retinae displayed some evidence of lamination, though this was reduced when compared to that in phenotypically wild-type siblings. Moreover, while cells populated the inner retina, and a GCL and INL were discernable, all *tet2*$^{-/-}$;*tet3*$^{-/-}$ mutants lacked a morphologically obvious optic nerve (Fig. 3.2K,L). Retinal defects remained prevalent at 4dpf with *tet2*$^{-/-}$;*tet3*$^{-/-}$ mutant retinae continuing to possess lamination defects and no apparent formation of an optic nerve, despite the presence of cells in the inner retinal area normally occupied by RGCs (Fig. 3.2N,O).

3.2.2 Cell cycle dynamics in early progenitor cells are disrupted in *tet2*$^{-/-}$;*tet3*$^{-/-}$

Because retinal cells in *tet2*$^{-/-}$;*tet3*$^{-/-}$ mutants appeared to be progenitor-like in morphology at 2dpf, and lamination defects persisted at 3dpf and 4dpf, a phenotype associated with elongated proliferation and/or defects in cell cycle exit in zebrafish (Cerveny et al., 2010; Uribe and Gross, 2010; Yamaguchi et al., 2005), we next determined cell cycle dynamics in *tet2*$^{-/-}$;*tet3*$^{-/-}$ mutant retinal cells. First we assayed the S to M phase progression by utilizing the percent labeled mitoses (PLM) assay (Locker et al., 2006; Quastler and Sherman, 1959; Uribe and Gross, 2010). Embryos were treated with a 15-minute bromodeoxyuridine (BrdU) pulse at 32hpf, washed, fixed at 30, 60, 90, and 120 minutes post-treatment, and immunostained for BrdU and phosphohistone H3 (pH3) (Fig. 3.5A). Cells in S-phase during the BrdU pulse (BrdU^{+}) and cells in late G2/M-phase at the time of fixation (pH3^{+}) were quantified. Cells that were proliferative during the BrdU pulse and then undergo mitosis are double positive (BrdU^{+},pH3^{+}). In contrast, cells that were not in S-phase during the BrdU pulse, but still undergo mitosis,

are only $pH3^+$. Thus, the proportion of these cells $\{(BrdU^+pH3^+)/pH3^+\}$ represent the 'labeled' mitotic events. In *tet2*$^{-/-}$;*tet3*$^{-/-}$ mutants, the percent labeled mitoses are significantly lower than sibling at all time points examined (Fig. 3.5B-D). Nearly 100% of 'labeled' RPCs in wild-type embryos completed the S to M phase transition by the end of 120-minute time window, while only 50% of *tet2*$^{-/-}$;*tet3*$^{-/-}$ mutants competed this transition. This indicates that between 32-34hpf, *tet2*$^{-/-}$;*tet3*$^{-/-}$ RPCs are progressing from S to M at a slower rate than wildtype RPCs.

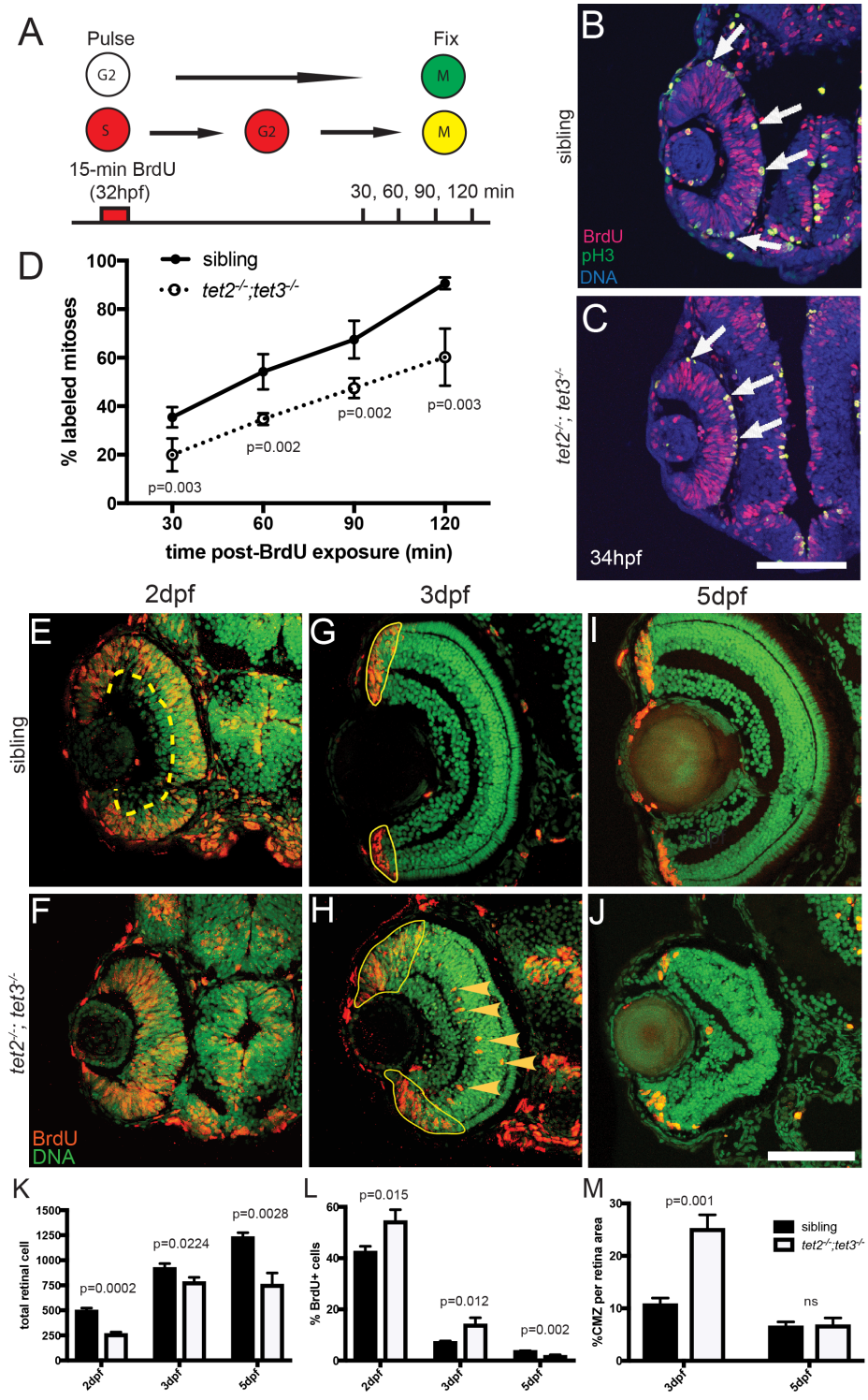

Figure 3.5: **RPC cell cycle dynamics are disrupted in *tet2*$^{-/-}$;*tet3*$^{-/-}$.**

(Figure 3.5 - continued) (A) Percent labeled mitoses (PLM) assay was performed by treating embryos for 15 minutes in bromodeoxyuridine (BrdU) pulse at 32hpf, rinsed, fixed at 30, 60, 90, and 120 minutes post-treatment for immunostaining. (B,C) Cells in S-phase during BrdU pulse ($BrdU^+$; red) and cells in G2/M-phase at fixation ($pH3^+$; green) were counted. Cells that were proliferative during the BrdU pulse and then undergo mitosis are double positive ($BrdU^+pH3^+$; yellow; arrows). (D) *tet2*$^{-/-}$*;tet3*$^{-/-}$ retinae show significantly lower proportion of labeled mitotic events ($[BrdU^+pH3^+]/pH3^+$) at all four time points. By the end of 120-minute window, nearly all 'labeled' RPCs completed S to M phase transition in wild-type siblings, compared to only 50% in *tet2*$^{-/-}$*;tet3*$^{-/-}$ mutants (n=5 embryos per condition per time point analyzed). (E-M) At later time points, BrdU incorporation assays over 2-hour time windows revealed that retinal progenitor cells remain proliferative longer and proliferate ectopically in *tet2*$^{-/-}$*;tet3*$^{-/-}$ mutants. (E,F) At 2dpf, cells within the central retina of wild-type sibling embryos (dotted line) are no longer proliferative, correlating with cell cycle exit and differentiation. (G,H) At 3dpf, the only proliferative cells in the wild-type retina are located in the CMZ (outlined). In *tet2*$^{-/-}$*;tet3*$^{-/-}$ mutants, this proliferative region is significantly expanded. Ectopically proliferating cells are also observed outside of the CMZ at significantly higher numbers than wild-type siblings (arrowheads). (I,J) At 5dpf, both sibling and *tet2*$^{-/-}$*;tet3*$^{-/-}$ eyes possess proliferative CMZs. (K) Total retinal cell count per section is significantly lower in *tet2*$^{-/-}$*;tet3*$^{-/-}$ mutant compared to sibling at all time points, correlating with microphthalmia. (L) *tet2*$^{-/-}$*;tet3*$^{-/-}$ mutant retinae possess a significantly higher percentage of proliferative ($BrdU^+$) cells at 2 and 3dpf, but lower at 5dpf. (M) Percentage of CMZ area per total retina is significantly higher in *tet2*$^{-/-}$*;tet3*$^{-/-}$ mutant at 3dpf, but not significantly different at 5dpf. N=3 embryos per condition per time point for E-M. Dorsal is up and anterior to the left in all images. All error bars = ± 1 S.D. All p-values calculated using two-tailed, unpaired t-test. Scale bar = 50μm in B-C; 80μm in E-J.

At later time points, BrdU incorporation assays over 2-hour pulse windows showed that, although *tet2*$^{-/-}$*;tet3*$^{-/-}$ mutant eyes were smaller and contained fewer cells (Fig. 3.5E-K), they contained a significantly higher percentage of BrdU^{+} cells at both 2dpf and 3dpf (Fig. 3.5L). At 2dpf, central retinal cells of wildtype embryos have exited the cell cycle and differentiated, while in *tet2*$^{-/-}$*;tet3*$^{-/-}$ mutant eyes the central retina remained proliferative (Fig. 3.5E,F). By 3dpf, cells of the wildtype retina have exited the cell cycle and differentiated except for those in the ciliary marginal zone (CMZ) at the retinal periphery which remains proliferative throughout the life of the animal (Fig. 3.5G,H) (Marcus et al., 1999). In *tet2*$^{-/-}$*;tet3*$^{-/-}$ mutants, most cells within the central retina were no longer BrdU^{+}, but the peripheral (CMZ) domain was significantly expanded (Fig. 3.5M). Moreover, ectopic proliferative cells were detected in the inner retina, outside of the *tet2*$^{-/-}$*;tet3*$^{-/-}$ CMZ (Fig. 3.5H), a phenotype not observed in wildtype siblings. The identity of these cells remains unclear, although we have ruled out the possibility of these being Müller glia (MG) cells, as they do not co-stain with the MG-specific marker, zrf-1/gfap (Fig. 3.5J). By 5dpf, a distinct zone of proliferation is present in *tet2*$^{-/-}$*;tet3*$^{-/-}$ mutants, indicating that the mutant RPCs eventually completed cell cycle exit, and that ectopically proliferative cells are no longer present (Fig 3.5I,J).

While the PLM and BrdU incorporation results suggest that elongation of the cell cycle during early retinal development underlies microphthalmia in *tet2*$^{-/-}$*;tet3*$^{-/-}$ mutants, apoptosis of RPCs or newly differentiated neurons could also contribute. To determine whether apoptosis plays a role in *tet2*$^{-/-}$*;tet3*$^{-/-}$ mutant retinal defects, we performed terminal deoxynucleotidyl transferase dUTP nick-end labeling (TUNEL) assays. No increase in apoptotic cells was detected in *tet2*$^{-/-}$*;tet3*$^{-/-}$ mutant until after 3dpf (Fig 3.6), indicating that apoptosis is unlikely to account for microphthalmia in *tet2*$^{-/-}$*;tet3*$^{-/-}$ mutants.

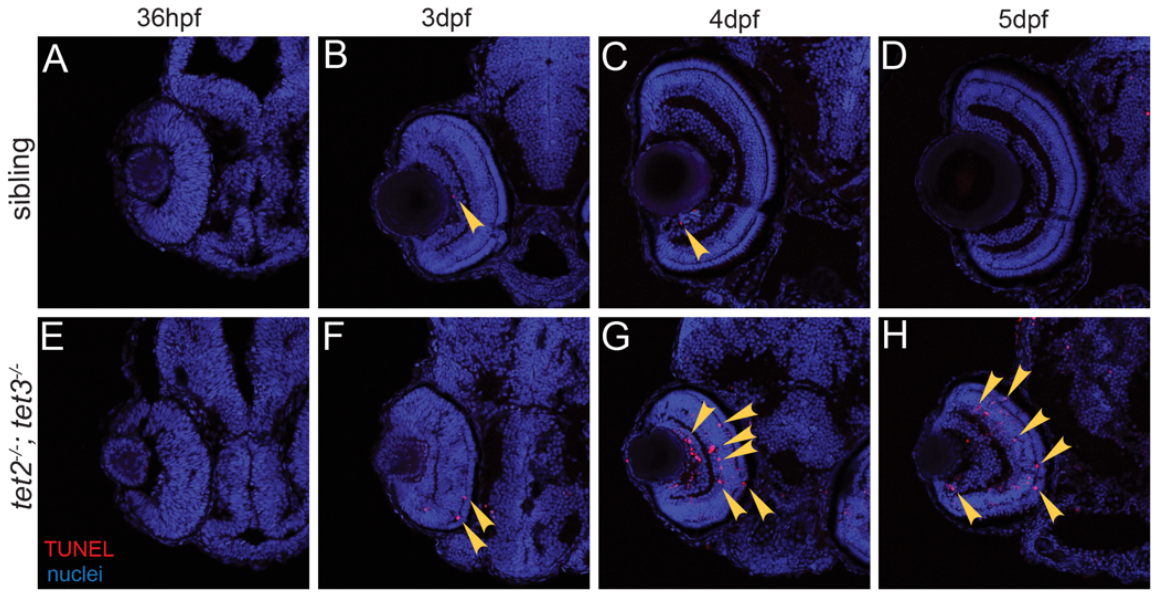

Figure 3.6: ***tet2*$^{-/-}$*;tet3*$^{-/-}$ embryos possesses few apoptotic cells prior to 3dpf.** TUNEL labeling was performed on cryosections of *tet2*$^{-/-}$*;tet3*$^{-/-}$ and sibling embryos at 36hpf, 3dpf, 4dpf, and 5dpf. No difference was observed at 36hpf (A,E), and few apoptotic cells are observed in *tet2*$^{-/-}$*;tet3*$^{-/-}$ at 3dpf (B,F; arrows). More apoptotic cells are observed in *tet2*$^{-/-}$*;tet3*$^{-/-}$ at 4dpf and 5dpf (C-D; G-H). Images are representatives of at least n=3 embryos examined. DNA (blue), TUNEL signal (red).

3.2.3 Retinal cell differentiation is impaired in *tet2*$^{-/-}$*;tet3*$^{-/-}$ embryos

Given the impaired retinal lamination and cell cycle progression defects in *tet2*$^{-/-}$*;tet3*$^{-/-}$ mutants, we next asked whether *tet2*$^{-/-}$*;tet3*$^{-/-}$ retinal cells differentiate into neurons and Müller glia. In wild-type embryos, all retinal neuron and glial cell types are differentiated by 72hpf [i.e. (He et al., 2012; Schmitt and Dowling, 1999; Uribe and Gross, 2010)]. An early neuronal marker, HuC/D, detects RGCs and amacrine cells (ACs) (Kay et al., 2001; Kim et al., 1996); in *tet2*$^{-/-}$*;tet3*$^{-/-}$ mutants, while HuC/D^{+} cells were detected in the GCL and INL, their number in the INL was significantly lower and the few HuC/D^{+} cells present were restricted to the central-most region of the retina (Fig. 3.7A,F; Fig. 3.8). *tet2*$^{-/-}$*;tet3*$^{-/-}$ mutants almost entirely lacked zpr1-expressing red/green

double cones (Fig. 3.7B,G), zpr-3-expressing rods (Fig. 3.7C,D, H,I), and zrf-1/gfap-expressing Müller glia (Fig. 3.7E,J). *tet2*$^{-/-}$*;tet3*$^{-/-}$ mutants possess positive immunoreactivity to Zn8/neurolin, which detects differentiated RGCs (Trevarrow and Kimmel, 1990), although the region of Zn8^{+} RGCs does not extend as far to the retinal periphery as in wild-type siblings (Fig. 3.7K,P). For each of the markers tested, immunoreactive cells were restricted to the central-most retina and none were detected more peripherally, corresponding with the expanded zone of proliferation detected at 3dpf in BrdU assays (Fig. 3.5).

In sections of *tet2*$^{-/-}$*;tet3*$^{-/-}$ mutant eyes, no optic nerve was detected (Figs. 3.2, 3.5, 3.6). It was formally possible that an optic nerve was present, but that axons were misrouted inside of the retina, a phenotype associated with defects in axonal pathfinding (Karlstrom et al., 1996). To exclude this possibility, we performed whole-mount chromogenic labeling using the Zn8/neurolin antibody to label RGC axons (Karlstrom et al., 1996). In 3dpf phenotypically wild-type siblings, Zn8 labeled intra-retinal RGC axons that extend along the vitreal surface of the eye to generate the optic nerve, as well as the optic nerve itself as it passes through the choroid fissure (CF) and into the optic chiasm (Fig. 3.7L). In *tet2*$^{-/-}$*;tet3*$^{-/-}$ mutants, Zn8 was only detected within the GCL and in a few axons within the choroid fissure (Fig. 3.7Q). As a more sensitive optic nerve labeling assay, we utilized a transgenic reporter, *Tg(isl2b:GFP)*^{zf7}, which expresses GFP in RGCs, and a subset of developing PRs (Pittman et al., 2008). Sibling embryos possessed a strong GFP signal in RGC cell bodies and axons along the entire length of the optic nerve, clearly visible in both section and whole-mount preparations (Fig. 3.7M-O). In *tet2*$^{-/-}$*;tet3*$^{-/-}$ mutants, while 12% of 41 mutant embryos lacked an optic nerve entirely, the remainder possessed an optic nerve, formed either bilaterally (39%) or unilaterally (49%); however, in these embryos, the optic nerve was thin and appeared to be composed of very few axons (Fig. 3.7M,R). For the few axons that did leave the eye, retinotectal projections appeared to be normal, even in embryos with unilateral optic nerves (Fig. 3.7R). Taken together, these data indicate that the terminal differentiation and morphogenesis of RGCs is affected in *tet2*$^{-/-}$*;tet3*$^{-/-}$ mutants, and not axon pathfinding.

Photoreceptors (PRs) undergo terminal differentiation and morphogenesis to form outer segments rich in photoreceptive molecules for phototransduction; this starts in the ventro-nasal patch and spreads through the retina (Fig. 3.7B-D) (Brzezinski and Reh, 2015; Schmitt and Dowling, 1996; Schmitt and Dowling, 1999). Sporadic rods and cones are detected in the central region of the retina in *tet2*$^{-/-}$*;tet3*$^{-/-}$ mutants, likely the earliest born PRs, but none are detected more peripherally (Fig. 3.7G-I). Moreover, in sibling embryos, while outer segments are well formed by 72hpf and highly immuoreactive to zpr-1 (arrestin3a) and zpr-3 (rhodopsin), the few PRs that differentiate in *tet2*$^{-/-}$*;tet3*$^{-/-}$ mutants possess little to no outer segment material (Fig. 3.7D,I). This is further highlighted by staining *isl2b:*GFP embryos with zpr-3. Only a few *isl2b:*GFP$^{-/-}$ PRs express zpr-3 and of those that do, they have nearly undetectable outer segments (Fig. 3.7O,T). Given the dramatic reductions in optic nerve size and in PR outer segment formation, we conclude that RGCs and PRs of *tet2*$^{-/-}$*;tet3*$^{-/-}$ mutants, despite expressing some markers of terminal differentiation, do not complete morphogenesis.

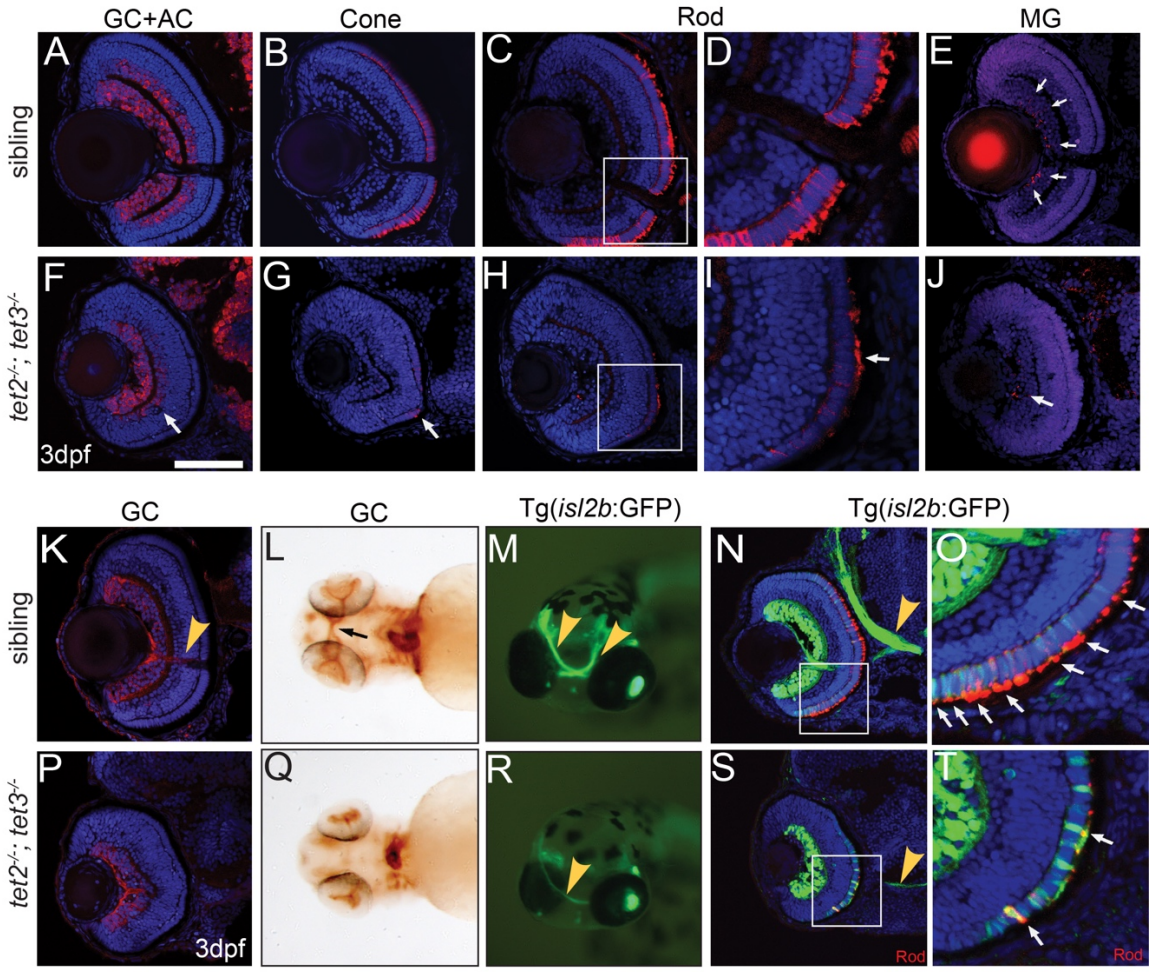

Figure 3.7: ***tet2*$^{-/-}$*;tet3*$^{-/-}$ retinal cells do not undergo terminal differentiation.**

(**Figure 3.7 - continued)** (A,F) HuC/D labels RGCs and amacrine cells (ACs), which are reduced in number and located only in the central region of the INL in *tet2*$^{-/-}$*;tet3*$^{-/-}$ retinae (arrow). (B,G) *tet2*$^{-/-}$*;tet3*$^{-/-}$ mutant retinae almost entirely lack zpr-1^{+} red/green cones (arrow); (C,D,H,I) possess few zpr-3^{+} rods (arrow), and of those that are zpr-3^{+}, outer segments are severely attenuated or almost absent (arrow). (E,J) *tet2*$^{-/-}$*;tet3*$^{-/-}$ retinae also possess few zrf-1^{+} Müller glia (arrows in wild-type). In all cases, marker^{+} cells are located in the central/ventral part of the retina. (K,P) Zn8 detects neurolin, a protein enriched on RGCs and the optic nerve (arrow in K). (L,Q) Zn8 staining reveals the optic nerve in the choroid fissure and optic chiasm (arrow) of wild-type embryos but not in *tet2*$^{-/-}$*;tet3*$^{-/-}$ mutants. (M-N) *isl2b*:GFP transgenics express GFP in RGCs and PRs, clearly labeling the optic nerve in whole-mount and section views (arrowhead). (R,S) The *tet2*$^{-/-}$*;tet3*$^{-/-}$ optic nerve is very thin, often unilaterally formed, but, when present, correctly routed to the brain. (O,T) The *isl2b*:GFP signal overlaps zpr-3 (rod) marker in the cell body and outer segments in siblings (arrows). In *tet2*$^{-/-}$*;tet3*$^{-/-}$, few *isl2b*:GFP+ cells are zpr-3+ (arrows), further suggesting that specified cells are not terminally differentiated. Few outer segments have also formed in *tet2*$^{-/-}$*;tet3*$^{-/-}$ mutants. DNA (blue), antibody stain (red). All images are 3dpf. n>5 for each marker. Dorsal is up and anterior to the left. Scale bar = 80μm in A-K, P.

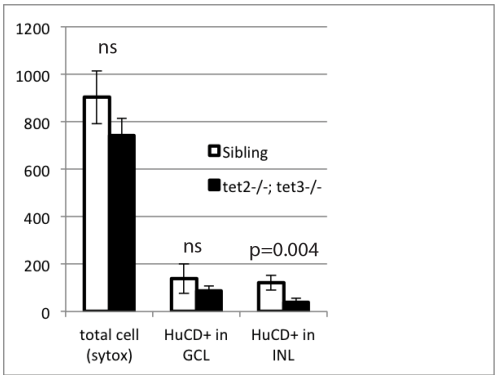

Figure 3.8: ***tet2*$^{-/-}$*;tet3*$^{-/-}$ embryos possesses fewer amacrine cells at 3dpf.** Number of HuC/D-positive neurons in the INL (amacrine cells) is significantly lower in *tet2*$^{-/-}$*;tet3*$^{-/-}$ eyes than in sibling, although the number of HuC/D-positive cells in the GCL (consisting of ganglion and displaced amacrine cells) is not significantly different. Error bars = ± 1 S.D. Significance cut-off for p-value = 0.05 (two-tailed, unpaired t-test).

3.2.4 Neuronal specification occurs normally in *tet2*$^{-/-}$;*tet3*$^{-/-}$ mutants

To begin to determine the molecular mechanism responsible for the reduction of terminally differentiated retinal neurons in *tet2*$^{-/-}$;*tet3*$^{-/-}$ mutants, we next asked if neuronal specification factors were properly expressed (Fig. 3.9). At 36hpf, *vsx2* is expressed in proliferative RPCs, and turned off as these cells exit the cell cycle and begin to differentiate (Fig. 3.9A, dotted area). In *tet2*$^{-/-}$;*tet3*$^{-/-}$ mutants, this zone of differentiation is noticeably smaller (Fig. 3.9F). At 48hpf, *vsx2* expression is localized in proliferative cells at the periphery of the retina (Fig. 3.9K) (Vitorino et al., 2009). This zone of *vsx2* expression was expanded in *tet2*$^{-/-}$;*tet3*$^{-/-}$ mutants when compared to wild-type siblings (Fig. 3.9P), corresponding to the expanded zone of proliferation observed in the BrdU labeling assay (Fig 3.5).

pax6a is normally expressed in RPCs, in addition to RGCs and ACs (Klimova and Kozmik, 2014); *neurod4* is expressed in ACs, horizontal and bipolar cells (Wang et al., 2003); *atoh7* is expressed in the committed precursor undergoing specification to become RGCs (Kay et al., 2001), and *crx* is expressed in specified PRs (Liu et al., 2001) (Fig. 3.9B-E, L-O). Despite terminal differentiation defects observed in retinae of *tet2*$^{-/-}$;*tet3*$^{-/-}$ mutants, they retain relatively normal spatial and temporal expression of specification markers at 36hpf and 48hpf (Fig. 3.9G-J, Q-T, respectively), suggesting that RPC specification is unaffected during retinal neurogenesis in the absence of tet2 and tet3 function.

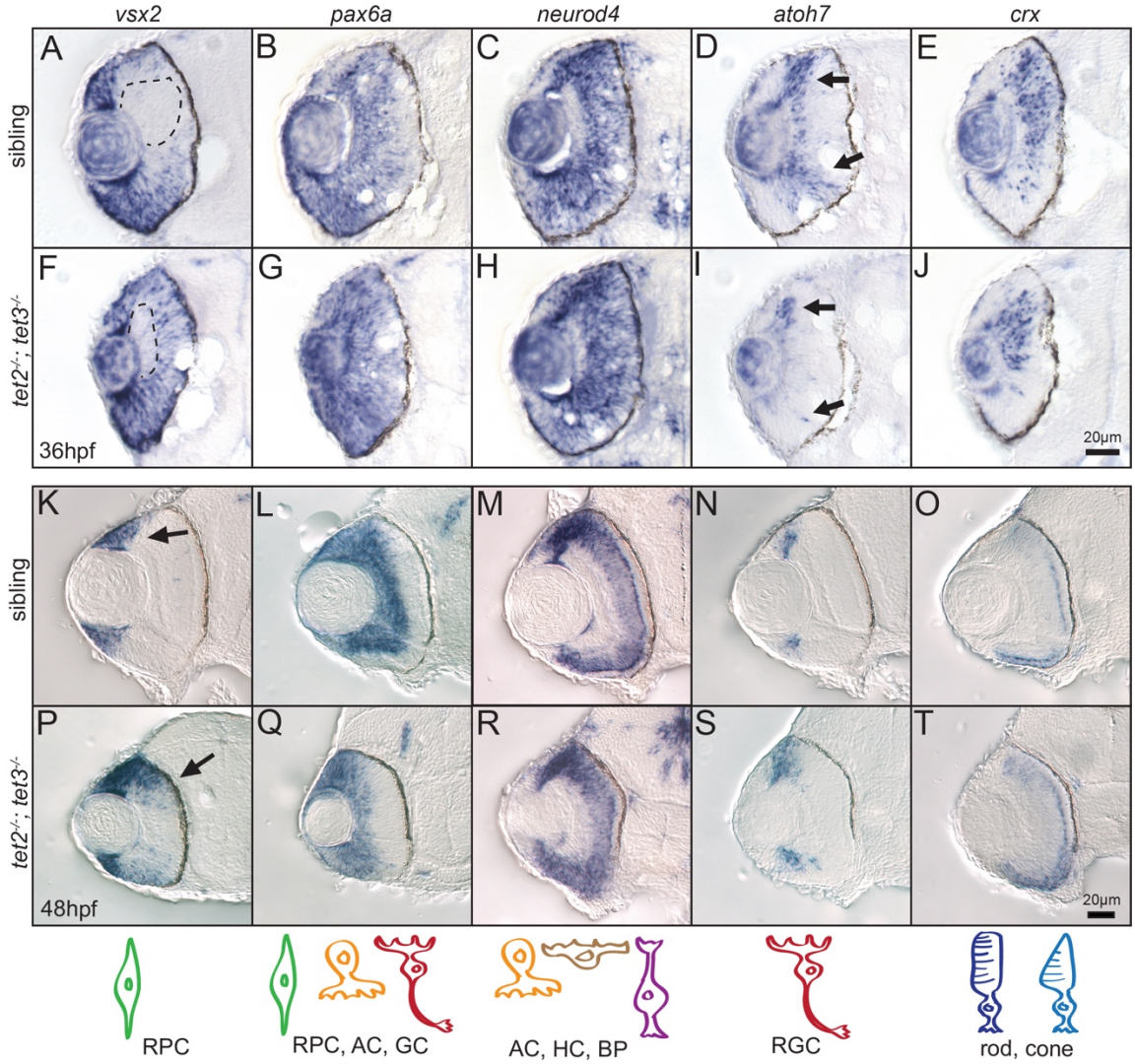

Figure 3.9: **Expression of retinal cell fate specification markers is largely normal in *tet2*$^{-/-}$;*tet3*$^{-/-}$ embryos.**

(**Figure 3.9 – continued)** Expression of genes involved in retinal neurogenesis was detected by in situ hybridization at 36hpf and 48hpf. (A,F) At 36hpf, *vsx2* is expressed in the retinal progenitor cells (RPCs) and turned off as they differentiate, first in the central retina (A, dotted area). This zone of differentiation is still present in *tet2*$^{-/-}$*;tet3*$^{-/-}$ mutant (F, dotted area). (K,P) At 48hpf, expression of *vsx2* is slightly expanded in *tet2*$^{-/-}$*;tet3*$^{-/-}$ mutants (K,P arrows) but otherwise appears in a normal pattern. (B,G,L,Q) Expression of *pax6a* (marker for RPC, amacrine, and ganglion cells) and *neuroD4* (marker for amacrine, horizontal, and bipolar cells) is relatively normal in *tet2*$^{-/-}$*;tet3*$^{-/-}$ mutants when compared to sibling embryos at both 36hpf and 48hpf. (D,I,N,S) Ganglion cell precursors express *atoh7* as they exit the RPC pool, become specified and start undergo differentiation. *atoh7* expression is present in the correct location in *tet2*$^{-/-}$*;tet3*$^{-/-}$ mutants, compared to sibling embryos at both 36hpf and 48hpf. (E,J,O,T) *crx* is expressed in cells fated to become photoreceptors (rods and cones) and this expression pattern appears relatively normal in *tet2*$^{-/-}$*;tet3*$^{-/-}$ mutants. Scale bar = 20μm. n>8 per gene for each time point. Drawings on bottom row represent cell type detected in each corresponding column.

3.2.5 tet2 and tet3 regulate cell non-autonomous effects during retinal neuron differentiation

Tet proteins are known to regulate both intrinsic and extrinsic pathways during development and differentiation events in a range of tissues. For example, tet activity is required intrinsically during hematopoiesis (Madzo et al., 2014) and B-cell differentiation (Orlanski et al., 2016). Recent studies have shown that Tet activity also modulates extrinsic pathways during development. In mouse embryos and embryonic stem cells, Tets function to negatively regulate Wnt pathway activity to balance mesoderm vs neuroectoderm fate choices (Li et al., 2016), and during mouse gastrulation, they modulate Nodal pathway activity by controlling the expression of Lefty1, a nodal inhibitor (Dai et al., 2016). Retinal cell type specification and differentiation depend on a multitude of intrinsic and extrinsic pathways [reviewed in (Bassett and Wallace, 2012; Yang, 2004)], and given that Tet activity can modulate both intrinsic and extrinsic

pathways during development, we next sought to determine whether tet2 and tet3 activities were required cell autonomously or cell non-autonomously during retinal neurogenesis. Chimeric embryos were generated by blastomere transplantation (Carmany-Rampey and Moens, 2006) to generate embryos whose retinae were composed of clones of wild-type and *tet2*$^{-/-}$;*tet3*$^{-/-}$ mutant cells. Donor embryos were injected with fluorescent dextran and clones were transplanted from labeled donors into unlabeled hosts at shield stage, targeting specifically to the retinal field where cells later develop into the retina (Woo and Fraser, 1995). At 3dpf, host embryos were analyzed through a combination of HuC/D and zpr-3 staining, to detect differentiated RGCs, ACs and/or rods (Fig. 3.10A). Sibling to wild-type transplants yielded clones of cells that differentiated normally (Fig. 3.10B). *tet2*$^{-/-}$;*tet3*$^{-/-}$ mutant cells transplanted into genetically wildtype hosts also differentiated normally into retinal neurons including RGCs, ACs, and rods, and the regions of the retina containing mutant clones were also properly laminated (Fig. 3.10C). Genetically wildtype cells transplanted into *tet2*$^{-/-}$;*tet3*$^{-/-}$ mutant host retinae failed to undergo neurogenesis and remained undifferentiated (Fig. 3.10D). Thus, the wild-type retina was able to support normal neurogenesis of *tet2*$^{-/-}$;*tet3*$^{-/-}$ mutant cells, while wild-type cells in a *tet2*$^{-/-}$;*tet3*$^{-/-}$ mutant retina did not differentiate properly. Taken together, these data demonstrate that tet2 and tet3 activity regulates retinal neuron differentiation via cell non-autonomous pathways, potentially by modulating the expression or activity of cell-extrinsic signaling molecules.

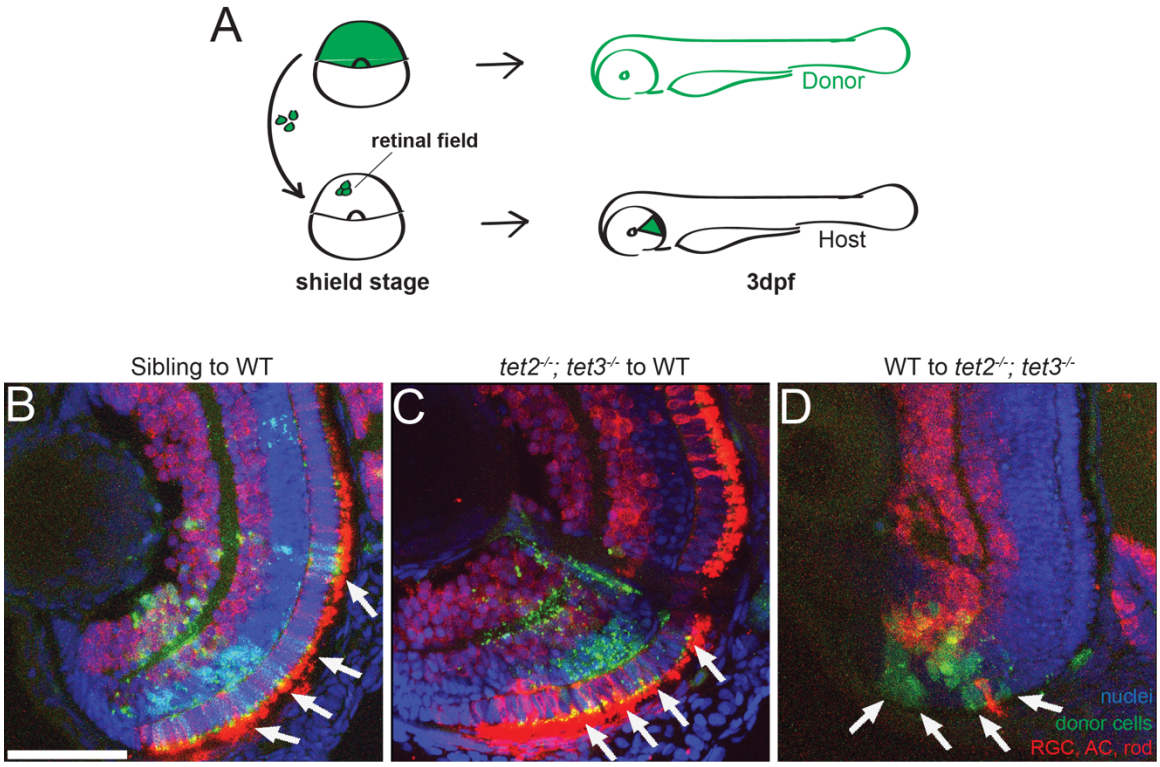

Figure 3.10: **Tet2 and tet3 regulate retinal neurogenesis cell non-autonomously.** (A) Chimeric embryos were generated by blastomere transplantation from fluorescent dextran-labeled donor into unlabeled host. Retinae composed of clones from wild-type and *tet2*$^{-/-}$;*tet3*$^{-/-}$ mutant cells were analyzed at 3dpf. (B) Sibling donor to wildtype host transplants yielded clones of donor cells (green; arrows) that differentiated normally into retinal neurons including AC, RGC, rods (red), and display proper lamination (n=5). (C) Similarly, *tet2*$^{-/-}$;*tet3*$^{-/-}$ mutant cells transplanted into genetically wildtype hosts also differentiated normally, and the regions of the retina containing mutant clones were also properly laminated (n=3). (D). Genetically wild-type donor cells transplanted into *tet2*$^{-/-}$;*tet3*$^{-/-}$ mutant host retinae failed to undergo neurogenesis and remained undifferentiated (green, arrows) (n=3). DNA (blue), HuC/D + zpr-3 antibody stain (red), transplanted donor clones (green). All images are 3dpf. Dorsal is up and anterior to the left. Scale bar = 50μm.

3.2.6 The Notch and Wnt pathways function downstream of tet2 and tet3

Because tet2 and tet3 appear to regulate retinal cell differentiation in a cell non-autonomous fashion, we next sought to identify potential factors responsible for these effects, utilizing both a candidate gene approach and an unbiased transcriptomic analysis. The Notch and Wnt pathways regulate retinal neurogenesis in mouse, zebrafish and Xenopus (Chiodini et al., 2013; Kubo et al., 2005; Uribe et al., 2012; Yaron et al., 2006); upregulation of these pathways prevents neuronal differentiation in a variety of contexts and resultant phenotypes resemble those in *tet2*$^{-/-}$*;tet3*$^{-/-}$ mutants (Bernardos et al., 2005; Meyers et al., 2012; Yamaguchi et al., 2005; Yaron et al., 2006). Thus, to determine whether the expression of Notch and Wnt pathway components, or the overall activities of the pathways, are regulated by tet2 and tet3, we first performed in situ hybridizations using probes specific to candidate genes in each pathway. In wild-type sibling embryos at 36hpf, *notch1a*, *deltaA*, and *ascl1a* are expressed in the eye but excluded from the inner part of central retina where RGCs, ACs and PRs are differentiating (Fig. 3.11A-C). However, in *tet2*$^{-/-}$*;tet3*$^{-/-}$ mutants, these genes are expressed uniformly throughout the retina, without a clear 'zone' of differentiation (Fig. 3.11D-F). Similarly, *lef1*, a downstream readout of the Wnt pathway (Borday et al., 2012), is expressed in the peripheral part of the retina of wild-type embryos at 36hpf (Fig. 3.11I), and this zone of expression is expanded in *tet2*$^{-/-}$*;tet3*$^{-/-}$ mutants (Fig. 3.11L). These expression patterns suggest that the Notch and Wnt pathways could have elevated activity in the *tet2*$^{-/-}$*;tet3*$^{-/-}$ mutant retina.

To complement these candidate gene studies, we performed RNA-seq using dissected *tet2*$^{-/-}$*;tet3*$^{-/-}$ mutant and phenotypically wildtype sibling eyes at 36hpf. Approximately 450 million reads were generated and mapped to GRCz10 (Trapnell et al., 2010; Trapnell et al., 2012) at 85.1% mapping efficiency. In *tet2*$^{-/-}$*;tet3*$^{-/-}$ mutants, 278 genes were downregulated and 489 genes were upregulated ($\log_2$ fold-change above 2) as compared to wild-type siblings (Supplemental Table 1 in (Seritrakul and Gross, 2017)). Gene ontology (GO) analysis for biological pathways categorized many of the

differentially expressed genes in the development of the visual perception, GPCR and cytokine-mediated signaling pathways, and ion transport (Fig. 3.11M,N). Notably, the highest upregulated gene was *wnt9b* ($\log_2$ fold-change=9.89), and many other members of the Wnt family were also upregulated (*wnt1, wnt3, wnt11r, wnt10a*) in *tet2*$^{-/-}$*;tet3*$^{-/-}$ mutant eyes. In situ hybridization using antisense probes for *wnt1* and *wnt9b* revealed that at 36hpf, while both genes showed faint expression at the peripheral edge of the retina in sibling embryos (Fig. 3.11G,H, n=7 for *wnt1*, n=10 for *wnt9b*), in *tet2*$^{-/-}$*;tet3*$^{-/-}$ mutants, both genes were expressed in expanded domains (Fig. 6J,K, n=7 for *wnt1*, n=8 for *wnt9b*), and at higher relative intensities, consistent with the RNAseq data. When combined with mosaic analyses, these data support a model in which extrinsic signals, likely including Notch and Wnt-related pathways, are regulated by tet2 and tet3 activity during retinal development and in their absence, these pathways are overactive and impair terminal differentiation of retinal neurons.

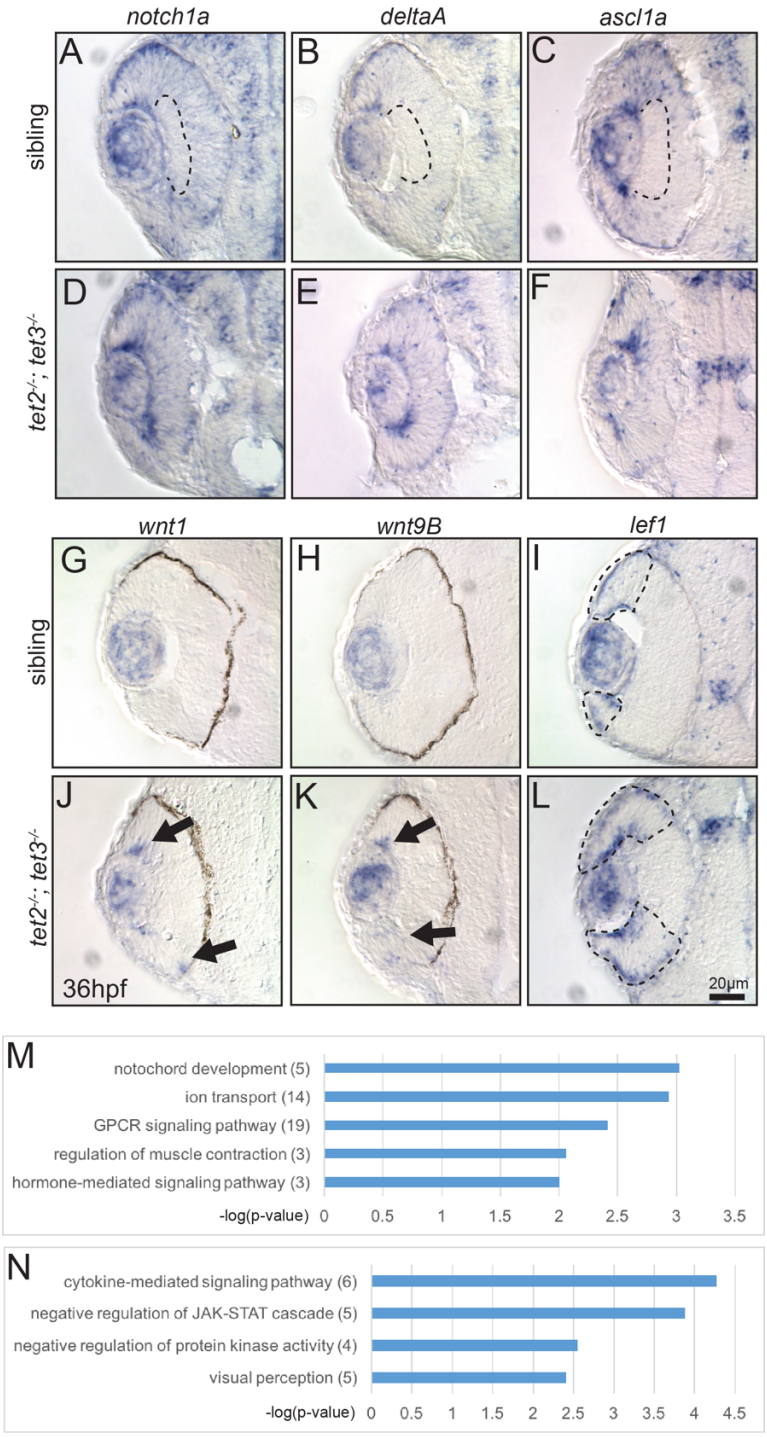

Figure 3.11: **Gene expression is altered in *tet2*$^{-/-}$;*tet3*$^{-/-}$ mutants at 36hpf and differentially expressed genes include those encoding components of the Notch and Wnt pathways.**

(**Figure 3.11 – continued**) (A-C) In sibling embryos at 36hpf, transcripts encoding components of Notch pathway (*notch1a*, *deltaA*, and *ascl1a*) are expressed in the eye but excluded from the inner part of central retina where cells have exited the cell cycle and differentiated (dotted area). (D-F) In *tet2*$^{-/-}$*;tet3*$^{-/-}$ mutants, these genes are expressed throughout the retina without a clear 'zone' of differentiation (n>8). (G,H,J,K) The expression domains of *wnt1* and *wnt9B*, are expanded in *tet2*$^{-/-}$*;tet3*$^{-/-}$ mutants (arrows; n=7 for *wnt1*, n=8 for *wnt9b*), consistent with RNA-Seq data. (I,L) Similarly, *lef1*, a downstream readout of Wnt pathway activity, is normally expressed in the peripheral edge of the retina, and this zone of expression is expanded in *tet2*$^{-/-}$*;tet3*$^{-/-}$ mutants (dotted areas; n>8). (M-N) Gene ontology (GO) analysis for biological pathways was performed using DAVID. Numbers in parentheses indicate number of genes enriched in each pathway. P-value cutoff = 0.01

To further test this model, we utilized pharmacological manipulations to determine if blocking Notch or Wnt activity could restore retinal neuron differentiation and morphogenesis in *tet2*$^{-/-}$*;tet3*$^{-/-}$ mutants. DAPT is a γ–secretase inhibitor that blocks the proteolytic cleavage of the intracellular domain of Notch, thus blocking its downstream signaling (Geling et al., 2002), and has been used extensively in zebrafish (Dovey et al., 2001; Geling et al., 2002; Kubo et al., 2005; Uribe et al., 2012). Inhibitor of Wnt Response (IWR-1-endo) stabilizes Axin2 and promotes β-catenin degradation, effectively inhibiting Wnt signaling (Chen et al., 2009). We exposed *tet2*$^{-/-}$*;tet3*$^{-/-}$ mutant and sibling embryos carrying the *isl2b*:GFP reporter to 50μM DAPT or 5μM IWR from the onset of neurogenesis (24hpf) until fixation at 3dpf. Interestingly, *tet2*$^{-/-}$*;tet3*$^{-/-}$ mutants treated with DAPT showed a significant increase in both the percentage of *isl2b*:GFP^{+} RGCs in the retina, and in optic nerve diameter, when compared to DMSO-treated controls (p=0.0044 and p=0.0010, respectively, 2-way ANOVA, Fig. 3.12A-H). No significant increases were detected in wild-type sibling embryos treated with either vehicle or DAPT (Fig. 3.12G-H). Similarly, *tet2*$^{-/-}$*;tet3*$^{-/-}$ embryos treated with IWR also showed a significant increase in percentage of *isl2b*:GFP^{+} RGCs and optic nerve diameter (p=0.0039 and p=0.0002, respectively, 2-way ANOVA, Fig. 3.12A-H). Neither treatment rescued expression of *isl2b*:GFP^{+} PRs however, suggesting that other extrinsic pathways are likely

regulated by tet2 and tet3 activity. These data support a model in which overactive Notch and/or Wnt signaling are partially responsible for neuronal differentiation/morphogenesis defects in *tet2*$^{-/-}$;*tet3*$^{-/-}$ mutants. To further test this model, we upregulated Wnt signaling by treating wild-type embryos from 24hpf to 3dpf with 2μM BIO, a GSK3β inhibitor (Borday et al., 2012; Nishiya et al., 2014). BIO-treated embryos showed dramatically reduced lamination, and decreases in the percentage of *isl2b*:GFP^{+} RGCs, and in optic nerve diameter relative to DMSO-treated controls, phenotypes that recapitulate those in *tet2*$^{-/-}$;*tet3*$^{-/-}$ mutants ($p=0.0024$ and $p<0.0001$, Fig. 3.12I-L). Taken together, these data demonstrate that Wnt and Notch signaling pathways act downstream of tet2 and tet3 activity to regulate RGC differentiation and morphogenesis during zebrafish retinal neurogenesis.

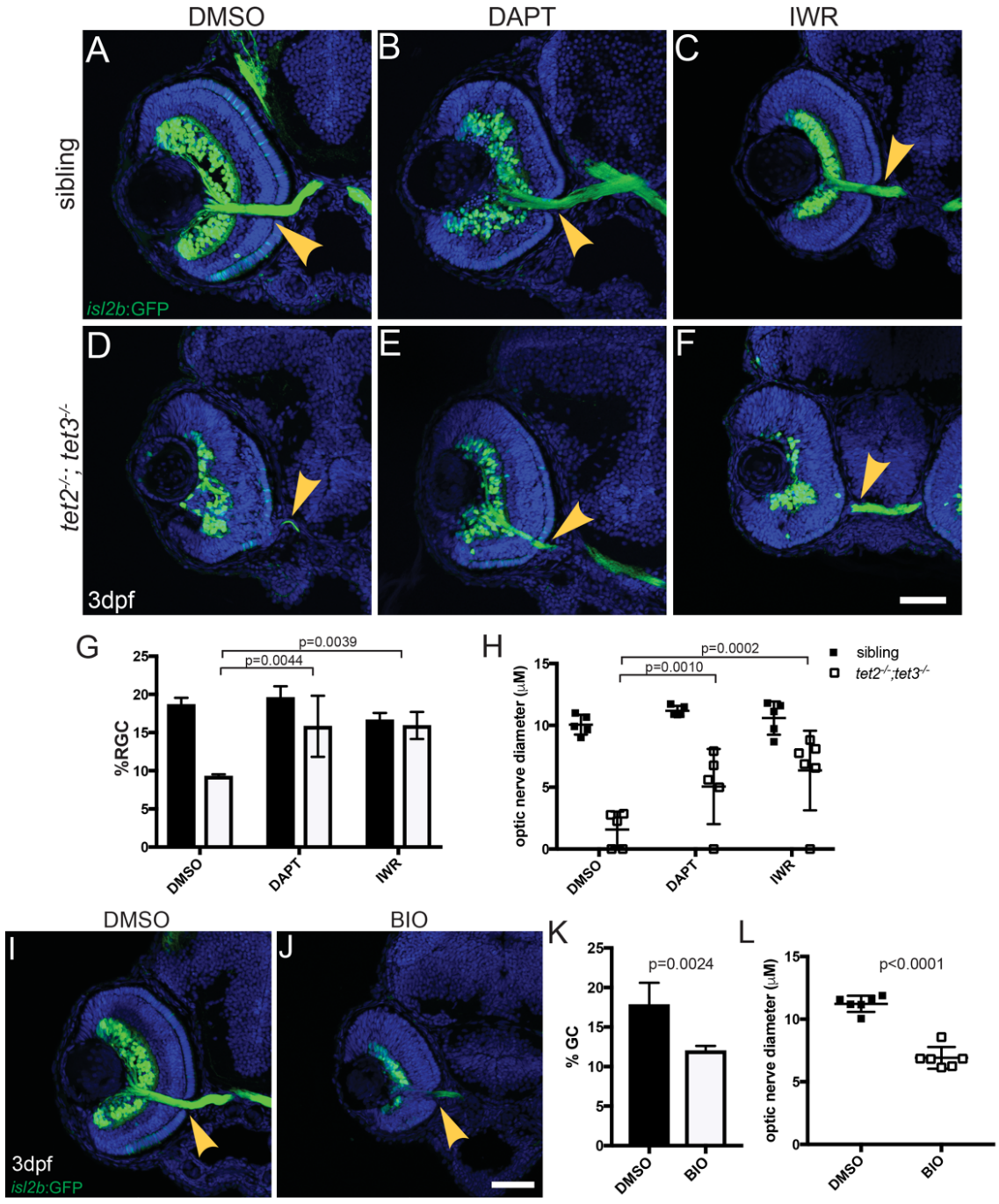

Figure 3.12: **Tet2 and Tet3 function upstream of the Notch and Wnt pathways during RGC differentiation and morphogenesis.**

(**Figure 3.12 - continued**) *tet2*$^{-/-}$*;tet3*$^{-/-}$ mutant and sibling embryos carrying *isl2b*:GFP were exposed to 50M DAPT, 5M IWR-1-endo, or 1% DMSO, during neurogenesis (24hpf to 72hpf) and analyzed for *isl2b*:GFP^{+} RGCs and axons. (A-C, G-H) No significant increases in optic nerve diameter or percentage of *isl2b*:GFP^{+} RGC per total cells (%RGC) were detected in sibling embryos treated with either 1% DMSO (vehicle), DAPT (Notch inhibitor), or IWR (Wnt inhibitor). (E,G,H) *tet2*$^{-/-}$*;tet3*$^{-/-}$ mutants treated with 50M DAPT showed a significant increase in both the percentage of *isl2b*:GFP^{+} RGCs per retina and optic nerve diameter, when compared to DMSO-treated controls. (F-H) *tet2*$^{-/-}$*;tet3*$^{-/-}$ embryos treated with 5M IWR also showed a significant increase in the percentage of *isl2b*:GFP^{+} RGCs per retina and optic nerve diameter. (I-L) Wnt signaling was upregulated by exposing wild-type embryos from 24hpf to 72hpf with 2M BIO, a GSK3 inhibitor. (I,J) BIO-treated wild-type embryos showed reduced lamination, (K) a decreased percentage of *isl2b*:GFP^{+} RGCs per retina, and (L) decreased optic nerve diameter relative to DMSO-treated controls, ($p=0.00236$ and $p<0.0001$). All error bars = ± 1 S.D.; n=5 embryos (A-H) and n=6 embryos (I-L) per condition analyzed; P-values calculated using two-way ANOVA with multiple comparisons (for G-H) and two-tailed, unpaired t-test (for K-L). Scale bar = 50μm

3.2.7 **Gene expression and 5hmC distribution is disrupted in *tet2*$^{-/-}$*;tet3*$^{-/-}$ mutants**

Despite early specification markers being expressed relatively normally during neurogenesis at 36 and 48hpf (Fig. 3.9), retinal neurons still do not terminally differentiate. Thus, we next sought to determine the gene expression signatures of retinal cells at 72hpf as a means to infer their identities, and we again utilized RNA-Seq to asses these. Similar to 36hpf assays, total RNA was extracted from dissected 72hpf *tet2*$^{-/-}$*;tet3*$^{-/-}$ eyes and used for RNA-seq. Sixty-six million reads were generated and mapped to GRCz10 at 88.7% mapping efficiency. Comparisons of wild-type to *tet2*$^{-/-}$*;tet3*$^{-/-}$ mutants resulted in the identification of 212 upregulated genes and 451 downregulated genes that passed a threshold of at least a 2 $\log_2$ fold-change (Fig 3.13A,B; Supplementary Table 2 in (Seritrakul and Gross, 2017)). GO analysis was performed and, as expected, downregulated genes in *tet2*$^{-/-}$*;tet3*$^{-/-}$ mutants are members of pathways tightly linked to differentiated PRs; i.e. GPCR signaling, membrane transport and ion transport. Indeed,

the most downregulated genes were those encoding proteins required for PR function such as opsins (e.g. *opn1mw1*, *opn1sw1*) and components of the visual cycle (e.g. *guca1c*, *guca1g, gnat2, grk1b*), an unsurprising result given the near absence of terminally differentiated PRs in *tet2*$^{-/-}$*;tet3*$^{-/-}$ mutants (Fig. 3.7). While the highest number of upregulated genes constituted the regulation of transcription category, other upregulated GO terms included skeletal muscle contraction and cardiac muscle contraction. These GO categories included genes that are not normally expressed in the eye; these included *nppa*, *vmhcl*, *chrng*, and *ucp1*, each of which encodes a protein involved in heart and muscle development and/or function (Fig 3.13B).

To verify RNAseq results, we performed in situ hybridization using probes specific to selected differentially expressed genes. Both *opn1mw1* and *opn1sw1* were expressed in PRs at 72hpf of wild-type sibling embryos, and expression was largely absent in *tet2*$^{-/-}$*;tet3*$^{-/-}$ mutants, except in a small patch of cells in the ventral retina that corresponds to the region where differentiated PRs are detected (Fig. 3.13C-H). *nppa* (*natriuretic peptide precursor a*) was the most highly upregulated gene in *tet2*$^{-/-}$*;tet3*$^{-/-}$ mutant eyes ($\log_2$ fold-change = 6.3). *nppa* encodes the precursor to a peptide required for cardiovascular function and is normally only expressed in the embryonic heart (Becker et al., 2014). All wild-type and *tet2*$^{-/-}$*;tet3*$^{-/-}$ mutant embryos showed normal heart expression of *nppa* (Fig. 3.13I). Interestingly, however, in *tet2*$^{-/-}$*;tet3*$^{-/-}$ mutants *nppa* was ectopically expressed in the retina and brain of all embryos (n=16/16 embryos), while no *nppa* expression was detected in the retina or brain of wild-type embryos (n=0/7 embryos) (Fig. 3.13I-K).

To gain molecular insight into the epigenetic regulation of the differentially expressed genes in *tet2*$^{-/-}$*;tet3*$^{-/-}$ embryos and whether these expression changes correlated with changes in 5mC or 5hmC deposition, we performed locus-specific methylation analyses, using bisulfite conversion followed by sequencing for 5mC + 5hmC, and a glucosylation-digestion assay for 5hmC specifically. We targeted regions surrounding the transcription start sites (TSSs) of *opn1sw1*, *opn1mw1*, a CpG island near a gene cluster that contains multiple opsins, including *opn1mw1*, as well as the TSS and gene body of

nppa (Table A3). Out of eleven targets selected for bisulfite sequencing, we observed no difference in methylation status at any CpG sites. In both wild-type and *tet2*$^{-/-}$;*tet3*$^{-/-}$ mutants, all TSSs and the *nppa* gene body were fully methylated, and the opsin cluster CpG island was fully unmethylated (Fig 3.13L and Table 3.1).

Bisulfite sequencing, while providing a nucleotide-resolution view of the methylation status of each CpG analyzed, is incapable of detecting differences between 5mC and 5hmC, meaning any CpG that appeared methylated in bisulfite assays could be either 5mC or 5hmC, or a mixture of both. To distinguish between these two epigenetic marks, we utilized glucosylation-digestion-based (Quest) assay to probe the presence of 5hmC at a glucosyl-sensitive restriction site, MspI. We selected targets that were located within or adjacent to the bisulfite-probed regions, due to the relatively sparse occurrence of MspI sites. Out of nine sites selected for analysis, only one, located in the gene body of *nppa*, showed a significant difference in 5hmC levels (Table 3.1). Interestingly, 5hmC was reduced to a nearly undetectable level at this site in *tet2*$^{-/-}$;*tet3*$^{-/-}$ mutants, when compared to WT siblings where ~25% of the residues were hydroxymethylated (Fig 3.13M). These data demonstrate that *tet2* and *tet3* mutations result in defects in 5mC to 5hmC conversion within the *nppa* gene body, which could contribute to misregulated expression of the locus.

Locus	Methylation status (Bisulfite sequencing)		5hmC enrichment (Quest)	
	Sibling	**tet2-/-;tet3-/-**	**Sibling**	**tet2-/-;tet3-/-**
opn1sw1 TSS site 1	methylated	methylated	no	no
opn1sw1 TSS site 2	methylated	methylated	no	no
opn1mw1 TSS site 1	methylated	methylated	*	*
opn1mw1 TSS site 2	methylated	methylated	*	*
opn1mw1 upstream MspI	methylated	methylated	no	no
opn1mw1-4 CpG island site 1	unmethylated	unmethylated	no	no
opn1mw1-4 CpG island site 2	unmethylated	unmethylated	no	no
nppa TSS site 1	methylated	methylated	no	no
nppa TSS site 2	methylated	methylated	no	no
nppa gene body site 1	methylated	methylated	yes	no
nppa gene body site 2	methylated	methylated	no	no

Table 3.1: Methylation status and 5hmC enrichment at candidate loci. * MspI site is absent from locus.

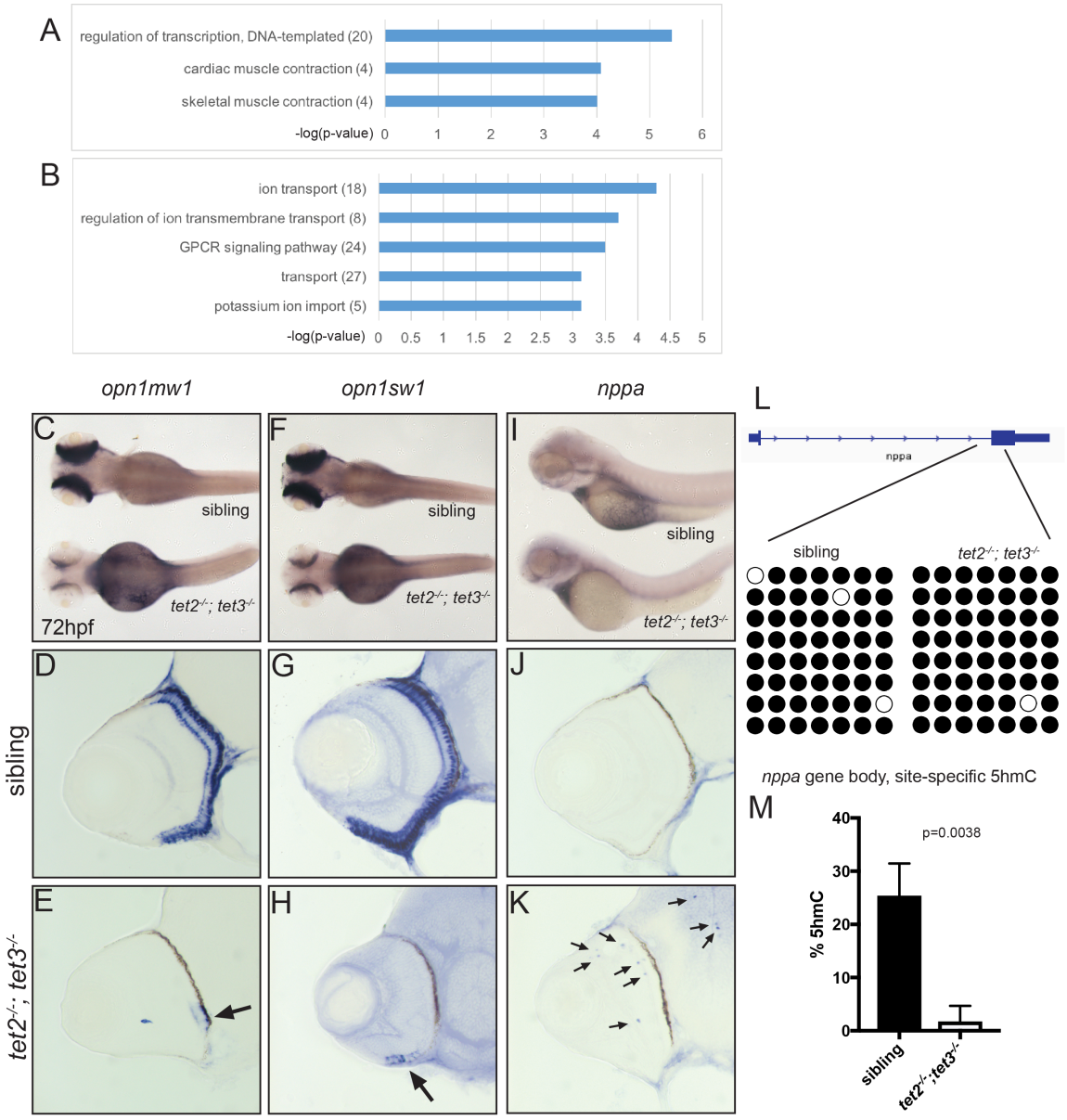

Figure 3.13: **Gene expression and 5hmC levels are abnormal in *tet2*$^{-/-}$;*tet3*$^{-/-}$ mutant eyes at 72hpf.**

(**Figure 3.13 – continued**) (A-B) GO analysis for biological pathways was performed using DAVID. Numbers in parentheses indicate number of genes enriched in each GO term. P-value cutoff = 0.001. In situ hybridization of (C-E) medium-wave sensitive (*opn1mw1*) and (F-H) short-wave sensitive opsin (*opn1sw1*). Transcripts of both genes are only detected in a few cells of the ventral retina in *tet2*$^{-/-}$;*tet3*$^{-/-}$ embryos (arrows in E,H; n>8). *natriuretic peptide a* (*nppa*) is normally expressed in the heart (I), and is not detected in the retina of wild-type embryos (J; n>8). (K) In *tet2*$^{-/-}$;*tet3*$^{-/-}$ embryos, ectopic *nppa* expressing cells were detected throughout the retina and brain (arrows) of all embryos examined (n=16/16). (L) Bisulfite sequencing did not identify any changes in DNA methylation in any of the sixteen RNAseq-identified target loci examined (Supplementary Table 4), including the *nppa* gene body. Bisulfite reads covering part of the first intron and second exon of *nppa* gene body are shown as black (methylated) or white (unmethylated). (M) Site-specific 5hmC quantification detected a significant, 20-fold reduction in 5hmC levels in the *nppa* gene body of 72hpf *tet2*$^{-/-}$;*tet3*$^{-/-}$ embryonic eye tissue, when compared to phenotypically wild-type siblings (p = 0.0038; two-tailed, unpaired t-test).

3.2.8 Overexpression of Tet resulted in retinal differentiation defect

Data from *tet2*$^{-/-}$;*tet3*$^{-/-}$ mutant analyses support a model in which site-specific Tet-mediated 5hmC formation is required for retinal neurogenesis and to regulate gene expression within the developing retina. To independently test this model, we generated inducible transgenic Tet overexpression lines to artificially drive 5mC → 5hmC conversion. To achieve this, we utilized the mouse Tet2 catalytic domain, which has been shown to ectopically generate 5hmC *in vitro*, along with its catalytically inactive mutant counterpart (Ito et al., 2010), and modified these by adding a T2A self-cleaving peptide (Kim et al., 2011) and mCherry and Myc tags, and placing them downstream of a *10xUAS* promoter. The resulting constructs were cloned into a zebrafish Tol2-compatible transgenic vector (Kwan et al., 2007) and multiple stable transgenic lines were created for each construct (*10xUAS*:nls-mCherry-t2a-myc-flag-Tet2WT^{au50}; *10xUAS*:nls-mCherry-t2a-myc-flag-Tet2mut^{au51}); only the strongest expressing lines, as evaluated by *cmlc2*:GFP reporter expression, were kept for analysis (Fig 3.14A,B).

To induce the transgene expression, adults carrying the Tet2 overexpression allele (Tet2-WT) or the catalytically inactive mutant allele (Tet2-mut) were crossed to *hsp70l*:Gal4^{vj005} heterozygotes (Gabher and Wittbrodt, 2004). Embryos were heat-shocked at 39.5°C at 18hpf for 30 minutes. We determined this timepoint based on the observation that the mCherry signal is first detectable ~6 hours after heatshock. Therefore, we timed this induction to coincide with the time at which RPCs are being specified and begin to differentiate (Vitorino et al., 2009). Embryos were sorted at 48-50hpf based on the expression of *cmlc2*:GFP from the promoter construct, and mCherry expression, indicating expression of the transgene (Fig 3.14C-J).

Heat-shocked embryos did not show any obvious morphological abnormalities (Fig. 3.14C-F), and they survived until at least 72hpf. To assess retinal development, embryos were cryosectioned and stained for the retinal cell type-specific markers. Tet2-WT-expressing embryos possessed relatively normal retinal lamination, and the differentiation of early-born cell types (RGCs and ACs) was relatively normal (Fig. 3.14E). However, most Tet2-WT embryos showed a distinct defect in the differentiation of later born rods and red/green cones (Fig 3.14K-R) where photoreceptors were absent in the retina except for small patches of cells on the ventral side (n=4/11 eyes) (Fig 3.14K-N). Zpr3 distribution was normal in Tet2-mut-expressing embryos (n=11/11 eyes) (Fig 3.14O-R). Taken together, these results indicate that, similar to loss of tet activity, ectopic or overactive Tet activity also disrupts the terminal differentiation of late born neurons during retinal development.

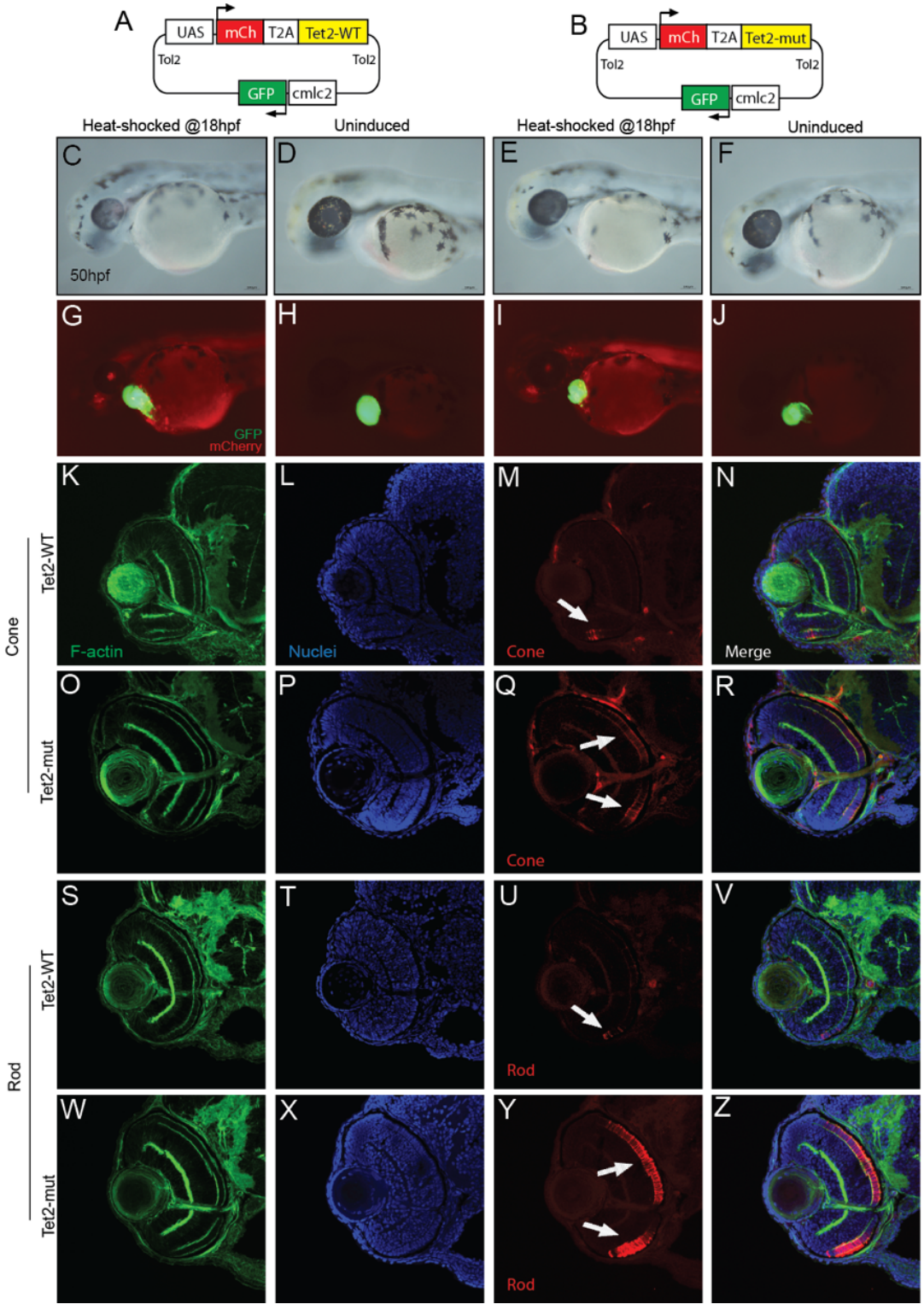

Figure 3.14: **Overexpression of Tet2 disrupts retinal neuron differentiation.**

Figure 3.14 – continued) Tol2 transgenic constructs carrying the mouse Tet2-WT domain (**A**) or a catalytically inactive Tet2-mut domain (**B**) were generated. *10xUAS*-drivenTet2 domains were separated by a T2A self-cleaving peptide and an mCherry reporter, and transgenic constructs carried a *cmlc2*:GFP reporter. Carriers were crossed to *hsp70l*:Gal4 drivers and embryos were heat-shocked at 18hpf for 30min at 39.5°C, and sorted at 48-50hpf (**C-F**). Induced embryos show ubiquitous mCherry expression and heart-specific GFP (**C,G**), while non-induced (non-heat-shocked) carriers express no detectable mCherry (**H,J**). Tet2-WT-expressing embryos are microphthalmic and lack differentiated rods (zpr3; red) when compared to Tet2-mut-expressing embryos (**Q,Y** arrows). F-Actin (green) green; DNA (cyan); Zpr3 (red). Dorsal up and anterior to the left in all images.

3.2 DISCUSSION

DNA hydroxymethylation and demethylation remain somewhat enigmatic processes in the field of epigenetics, with Tet protein function having only been identified recently [reviewed in (Pastor et al., 2013)]. Tet proteins are the main drivers of 5mC to 5hmC conversion and thereby key regulators of DNA demethylation (Tahiliani et al., 2009). However, in recent years, it has also become evident that they play roles in tissue-specific regulation of gene expression during development (Colquitt et al., 2013; Ge et al., 2014; Li et al., 2015; Madzo et al., 2014; Xu et al., 2012). Despite these studies, we still know little about their developmental functions, and we know virtually nothing about Tet function during eye development.

Here, we demonstrate that tet2 and tet3 play a critical role during development of the zebrafish retina. Our data indicate that tet2 and tet3 function redundantly in zebrafish to generate 5hmC, consistent with a recent report (Li et al., 2015). We demonstrate that tet2 and tet3 are critical regulators of retinal cell differentiation and morphogenesis, and that they act during early retinal development by modulating cell non-autonomous pathways. Loss of tet2 and tet3 function resulted in specific defects in retinogenesis, where the RPC population was transiently expanded. Despite relatively normal specification events, retinal cells failed to differentiate and, in the case of RGCs and PRs, failed to undergo terminal morphogenesis. These defects also correlated with mis-regulated gene expression and locus-specific defects in 5hmC formation in subsets of retinal cells at later stages of development. Based on these results, we propose a model wherein Tet proteins function to regulate gene expression during the differentiation of retinal cell types. In the absence of tet2 and tet3 function, gene expression is misregulated and differentiation and terminal morphogenesis of retinal neurons is perturbed (Fig. 3.15).

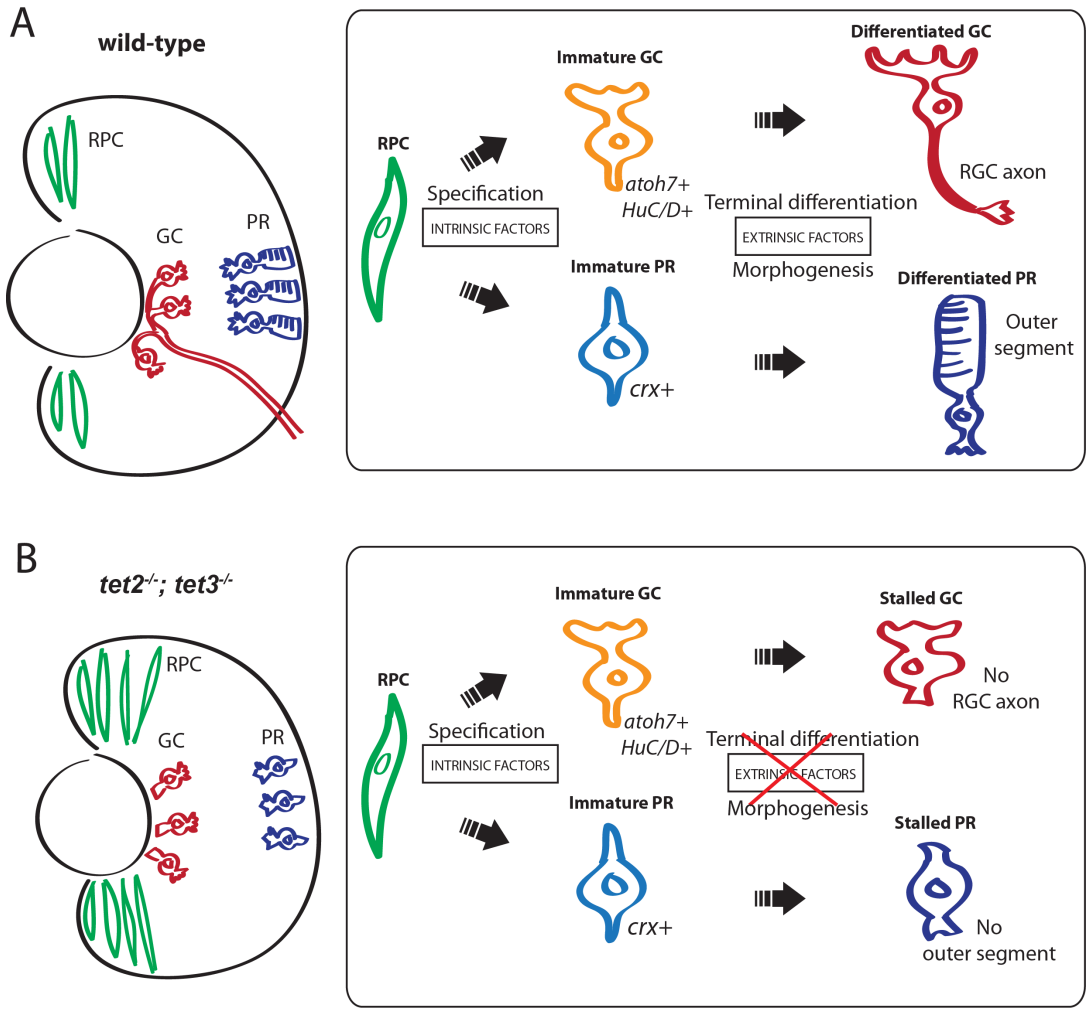

Figure 3.15: **Schematic of Tet function during zebrafish retinal development.** (A) During normal development, Tet proteins mediate epigenetic programming in retinal progenitor cells (RPCs; green) as they are specified into retinal neurons: retinal ganglion cells (RGCs, red) and photoreceptors (PRs, blue). These neurons undergo terminal differentiation and morphogenesis, generating axons that bundle into an optic nerve (for RGCs) and outer segments (for PRs). (B) In *tet2*$^{-/-}$;*tet3*$^{-/-}$ embryos, the RPC population is transiently expanded at early stages, but normal by 5dpf. Retinal neurons, despite being properly specified, fail to undergo terminal differentiation, and of those RGCs and PRs that do differentiate, many fail to undergo morphogenesis. These defects are likely caused by misregulation of cell-extrinsic factors required for terminal differentiation.

Similar to the cellular differentiation defects identified here in *tet2*$^{-/-}$*;tet3*$^{-/-}$ retinae, loss of Tet2 and/or Tet3 function results in differentiation defects in the hematopoietic system of zebrafish (Ge et al., 2014; Li et al., 2015) and humans (Madzo et al., 2014). During hematopoiesis, Tet proteins turn on genes involved in erythropoiesis by hydroxymethylating and/or demethylating the their promoters, thereby enabling expression and ultimately, triggering differentiate. Conversely, it is well established that Tet1 is a critical player in stem cell maintenance where it functions to inhibit differentiation potential (Etchegaray et al., 2015; Neri et al., 2015), and loss of Tet1 impairs ESC self-renewal (Ito et al., 2010). Therefore, Tet proteins play distinct roles in different contexts: the stimulate differentiation in tissue specific contexts (e.g. retinal, blood cells), and suppress differentiation in stem cells.

In *tet2*$^{-/-}$*;tet3*$^{-/-}$ mutants, early-born retinal cell types, RGCs and ACs, were less affected than later-born ones (cones, rods and Müller glia), which were almost completely absent. RGCs in *tet2*$^{-/-}$*;tet3*$^{-/-}$ mutants were *isl2b*:GFP^{+} and expressed Zn8, a marker of terminal differentiation (Trevarrow and Kimmel, 1990). However, Zn8^{+} RGCs were restricted to the central retina, and in a subset of mutants, no optic nerve (ON) was present, while in the remainder, a severely attenuated ON formed. In these latter embryos, we speculate that the RGCs generating axons are most likely the 'pioneer' axons (Pittman et al., 2008) that originate from the few early-born RGCs that undergo terminal differentiation and morphogenesis, while the majority of RGCs fail to complete morphogenesis to form an axon. ACs were detected in fewer number and located in an even more limited zone within the central retina of *tet2*$^{-/-}$*;tet3*$^{-/-}$ mutants, and the few differentiated red/green double cones, rods or Müller glia detected in *tet2*$^{-/-}$*;tet3*$^{-/-}$ mutant retinae were always located in the central retina. Specification and differentiation in the zebrafish retina initiates in the ventronasal patch, adjacent to the optic nerve, and proceeds in a central to peripheral gradient (Hu and Easter, 1999). That differentiated cells in *tet2*$^{-/-}$*;tet3*$^{-/-}$ mutants reside in these retinal regions strongly suggests that they represent the first born cells of each retinal cell type. Zebrafish embryos are also endowed with a maternally-derived supply of mRNA and protein (reviewed in (Langley

et al., 2014)). Therefore, it is possible that the centrally located and partially differentiated early born cell types in *tet2*$^{-/-}$;*tet3*$^{-/-}$ mutants reflect a "maternal rescue", and that defects in later born cell types result from depletion of maternally supplied tet2 and/or tet3. However, *tet* transcripts are not maternally deposited in zebrafish (Almeida et al., 2012b), and our 36hpf RNA-Seq analysis from *tet2*$^{-/-}$;*tet3*$^{-/-}$ mutants detects no expression of wild-type (maternally derived) *tet2* or *tet3* transcripts, making this scenario unlikely.

Alternatively, tet activity could become progressively more important in RPCs as they transition from producing early born cell types to later born ones, and as the specified cells undergo terminal differentiation and morphogenesis. Moreover, that the first born neurons of each class appeared to partially differentiate, tet2 and tet3 activity could become more important over time in each class of retinal neuron, where the first born neurons of the class develop independent of tet2 and tet3 function, while subsequent neurons require it. None-the-less, in this scenario, early born cells (and cell types) still require tet activity for terminal differentiation, because the majority of RGCs in *tet2*$^{-/-}$;*tet3*$^{-/-}$ mutants, though properly specified, do not form axons , and similarly in PRs, centrally-located cells are specified (*crx*$^{+}$) and begin to differentiate (*isl2b*:GFP^{+}, zpr-1^{+} or zpr-3^{+}), but do not complete outer segment morphogenesis. These data suggest that while RGCs and early born PRs may be refractory to the absence of tet activity during the earliest phases of differentiation, tet activity is still required for terminal morphogenesis. Our speculation that epigenetic regulation plays an important role in terminal differentiation of retinal neurons is also supported by recent evidence in mouse where disruption of Dnmt1, 3a and 3b resulted in severe retinal defects in which some PRs were present, but they appeared disorganized and failed to form outer segments (Singh et al., 2016), defects reminiscent to those in *tet2*$^{-/-}$;*tet3*$^{-/-}$ mutants.

Mosaic analyses reveal that *tet2*$^{-/-}$;*tet3*$^{-/-}$ retinal phenotype occurs cell non-autonomously, and thus, that the effects of tet2 and tet3 loss of function during early retinal development are mediated by cell extrinsic events. Through a combination of candidate gene assays, transcriptomics and pharmacological manipulations, we

demonstrate that elevated Notch and Wnt pathway activity is partially responsible for defects in retinal neurogenesis in *tet2*$^{-/-}$;*tet3*$^{-/-}$ mutants. However, because blocking these pathways only provided partial rescue of retinal defects in *tet2*$^{-/-}$;*tet3*$^{-/-}$ mutants, other signaling pathways are likely to be involved and modulated by tet activity during early retinal development. Hedgehog-PKA, TGFβ/BMP, and FGF are all cell-extrinsic pathways known to contribute to retinal neurogenesis, making these attractive candidates (Davis et al., 2000; Masai et al., 2005; Patel and McFarlane, 2000). While surprising, these cell non-autonomous results are consistent with those from several other recently published studies on Tet function. Indeed, Tet activity was demonstrated to modulate Nodal activity during mouse gastrulation by intrinsically regulating the methylation status and expression of the Nodal inhibitors, Lefty1 and Lefty2 (Dai et al., 2016). Tet activity has also recently been shown to modulate Wnt ligands or Wnt target gene activity in several contexts, either directly or indirectly (Kim et al., 2016; Li et al., 2016). Perhaps the most interesting of these recent studies showed that in mouse ESCs and early embryos, Tet activity is required to balance neuroectoderm vs. mesoderm fates and to inhibit Wnt signaling (Li et al., 2016). In Tet1/2/3 deficient ES cells and embryos, neural cell fates were lost and instead, mesodermal fates like cardiac muscle were detected. These fate changes correlated with increased promoter methylation and decreased expression of the Wnt inhibitor, Sfrp4 and hyperactive Wnt pathway activity. This parallels what we observe in the *tet2*$^{-/-}$;*tet3*$^{-/-}$ mutant zebrafish retina, and is exciting, because it suggests that Tet-mediated modulation of the Wnt pathway, and possibly other cell-extrinsic signaling pathways, may be a conserved function for Tets during embryonic development and organogenesis. Finally, Tet function may also influence chromatin accessibility at the genomic regions surrounding Notch and Wnt genes, enabling other transcriptional regulators to access these loci. Tets have been shown to function in regulating local chromatin environments (Hahn et al., 2013; Lio et al., 2016; Shen et al., 2013; Xu et al., 2011).

From RNA-seq analysis, we detected a suite of cardiac and muscle genes ectopically expressed in the retinae of 72hpf *tet2*$^{-/-}$;*tet3*$^{-/-}$ mutants. Further analysis of one

of these, *nppa*, revealed ectopic expression in the brain and eyes of *tet2*$^{-/-}$*;tet3*$^{-/-}$ mutants. This expression change correlated with an almost complete loss of gene body 5hmC deposition at the *nppa* locus. Locus-specific effects like this indicate that tet2 and tet3 may also function intrinsically during retinal development, in addition to modulating cell extrinsic pathways. Tet-mediated formation of 5hmC serves as a precursor to demethylation, or as an activating mark in its own right (Hahn et al., 2013). Tet-mediated 5hmC formation could play a direct role in silencing ectopic gene expression for genes like *nppa* during retinal development. Gene body 5mC methylation positively correlates with expression (Jin et al., 2012); therefore, in this scenario, Tet-mediated conversion to 5hmC likely serves as a precursor for subsequent demethylation and silencing. In *tet2*$^{-/-}$*;tet3*$^{-/-}$ mutants, 5mC is not converted to 5hmC, and the target locus (*nppa*) is ectopically expressed by retinal and brain cells. However, Tet-activity is also required for gene body 5hmC formation that is thought to serve as an active mark (Hahn et al., 2013). Thus, an alternative model can be envisioned wherein ectopic *nppa* expression also reflects indirect consequences of loss of tet2 and tet3 activity. In this model, an intermediate silencer/repressor gene is not properly expressed by *tet2*$^{-/-}$*;tet3*$^{-/-}$ mutant retinal cells due to the absence of activating 5hmC marks. This scenario is not unprecedented; Li et al recently demonstrated that *tet2*$^{-/-}$*;tet3*$^{-/-}$ mutants possess defects in hematopoiesis, but the mutants showed no changes in methylation or hydroxymethylation at key genes in the hematopoietic network; instead, in this context, Tet-mediated effects are likely the result of mis-regulation of the Notch pathway (Li et al., 2015). Similarly, mutation of the *de novo* DNA methyltransferase *dnmt3bb1* results in significantly altered expression of many hematopoietic and endothelial genes, although very few of these showed any changes in DNA methylation (Gore et al., 2016). Therefore, genes identified as differentially expressed in *tet2*$^{-/-}$*;tet3*$^{-/-}$ mutants that displayed no changes in 5mC or 5hmC may not be direct targets of tet2 and tet3, but rather, reflect complex intrinsic or extrinsic regulatory pathways modulated by tet2 and tet3 activity. Testing this prediction will require genome-wide profiling of 5mC and 5hmC in the eye over multiple developmental time points and correlating these data with gene expression in a gene-by-gene fashion, as

well as generating conditional loss of function tet2 and tet3 alleles such that their functions during later retinal development can be elucidated. CRISPR/Cas9 technology now makes this possible in zebrafish (Albadri et al., 2017).

It is known that CXXC4/IDAX, a protein with homology to the tet3 N-terminal domain, functions as a direct inhibitor of Wnt signaling by competitively binding with Axin to Dvl (Hino et al., 2001). In *tet2*$^{-/-}$*;tet3*$^{-/-}$ mutants, we detected no tet3 protein and 5hmC was almost completely absent from the genome, supporting a catalytic role for tet2 and tet3 during development. Significant changes in the expression of several Wnt ligands was detected in *tet2*$^{-/-}$*;tet3*$^{-/-}$ mutants, and these changes could result from this lack of catalytic activity (i.e. directly, from the lack of 5hmC formation at the loci, or indirectly, from the lack of 5hmC at loci encoding modulators of expression the Wnt ligands). However, alternatively, changes in Wnt ligand expression could also result from loss of the tet3 N-terminal CXXC domain, which functions independently of tet3 catalytic activity. Our data cannot yet differentiate between these possibilities. Importantly, this highlights the need to better understand the catalytic vs non-catalytic functions of tet proteins in specific tissues and organs, where tet proteins could regulate gene expression in several different ways.

Finally, in addition to DNA modifying enzymes like the Tets, chromatin regulators such as histone deacetylases and histone demethylases have also been shown to modulate cell extrinsic pathways during early retinal development (Lussi et al., 2016; Yamaguchi et al., 2005). When combined with our work, these studies highlight that the epigenetic regulation of signaling events during development is likely to be a more significant and complex layer of regulation than previously realized. Deciphering this complex epigenetic regulation will require a comprehensive, genome-wide approach encompassing multiple profiling strategies (e.g. bisulfite sequencing, oxidative bisulfite sequencing, RNA-seq, ChIP-seq, and ATAC-Seq) in pure, isolated RPCs and differentiated retinal cell types from both wild-type embryos, as well as embryos deficient in key enzymes operating in these epigenetic pathways, like tet2 and tet3.

Chapter 4: Methylome and Hydroxymethylome Analysis of Retinal Progenitor Cells (RPCs)

4.1 INTRODUCTION

Genome-wide epigenetic profiling of combined DNA methylation and hydroxymethylation in various cell types has been a challenging goal since the discovery less than a decade ago of the conversion of 5mC to 5hmC and Tet protein function (Tahiliani et al., 2009). The majority of earlier efforts in epigenomic profiling have been aimed at quantifying a combined (5mC+5hmC) mark using bisulfite conversion followed by sequencing (BS-seq) (Clark et al., 1994). These include approaches such as whole-genome shotgun bisulfite (WGBS) and reduced representation bisulfite (RRBS) approaches (Meissner et al., 2005), which are typically done in conjunction with other relevant chromatin modification and transcriptome profiling to obtain a comprehensive view of the genomic distribution of epigenetic marks and gene expression. For example, in human embryonic stem (ES) cells, a large scale profiling was done using WGBS, histone methylation ChIP-seq, and RNAseq (Gifford et al., 2013). This study identified epigenetic and gene expression changes in both pluripotent and developmental genes when human ES cells undergo lineage specification into germ layers. However, this was done using an *in vitro* directed differentiation paradigm and may not necessarily recapitulate the process *in vivo*.

WGBS has been utilized in multiple *in vivo* models and a common theme emerges across studies where DNA methylation patterns often inversely correlate with gene expression, and each cell population exhibits a distinct methylation pattern and gene expression profile. For example, WGBS revealed unique DNA methylation profiles of human cells and tissues such as sperm, oocyte, blastocyst, ES cells and blood (Okae et al., 2014). Additionally, subpopulations of cell types within the same tissue may also

exhibit distinct methylation patterns, which has been shown both in mammalian brain neuronal subtypes (Mo et al., 2015) and among rods and cone photoreceptor cells within the retina (Mo et al., 2016). Within the eye, genome-wide DNA methylation patterns have been profiled using WGBS from multiple time points during chicken retinal and corneal tissue development (Lee et al., 2017). Furthermore, a recent study revealed distinct chromatin states during development of both mouse and human retina using comprehensive profiling of DNA methylation and numerous chromatin markers (Aldiri et al., 2017). It is important to note that these methylation profiling strategies relied on WGBS and RRBS, both of which do not distinguish between 5mC and 5hmC, instead, both marks are read as "methylated" during bisulfite sequencing (Huang et al., 2010). Given the antagonistic nature of these two opposing epigenetic marks, it is necessary to generate separate genome-wide profiles where each mark is quantified independently. Distinguishing 5mC from 5hmC during profiling is inherently challenging due to the fact that both marks are chemically nearly identical except for a single hydroxyl group. Several methods have been developed over the past decade that enable simultaneous profiling of 5mC and 5hmC at the genome-wide level. These methods can be categorized into two major approaches: (1) affinity-enrichment profiling and (2) bisulfite-based profiling.

Affinity enrichment strategies rely on the ability of antibodies to bind uniquely to 5mC and 5hmC. Then, they are pulled down with the fragmented DNA to be sequenced (MeDIP-seq and hMeDIP-seq). A weakness to this approach is the potential for antibody affinity bias towards certain repeat-rich regions of the genome (Wu and Zhang, 2011). Alternatively, 5hmC can be converted into a glycosylated form (g5hmC) using beta-glucosyltransferase (βGT), which can be pulled down using a g5hmC-binding protein, JBP1, and sequenced (JBP1-seq). However, the affinity of JBP1 for its target has been

shown to be inferior to that of a 5hmC-specific antibody (Robertson et al., 2011). Additionally, a major limitation of affinity-based profiling strategies is their resolution; the output reads only reveal 5mC/5hmC within 'windows' throughout the genome and the finest resolution of these windows is limited by the size of the DNA fragments after shearing and not at single-nucleotide level.

Bisulfite-based 5hmC profiling methods have built upon the original bisulfite sequencing strategy (Clark et al., 1994), and include a modification step prior to sodium bisulfite treatment. Tet-assisted bisulfite sequencing (TAB-seq) utilizes βGT to convert 5hmC to g5hmC, which becomes "protected" and remains unconverted during bisulfite treatment, while 5mC is globally demethylated using a recombinant Tet1 enzyme and is fully converted to thymine (T) (Yu et al., 2012). This is often referred to as a *direct* profiling, because it reveals 5hmC levels directly. However, this strategy still relies on the enzymatic activity and accessibility of Tet1 to all 5mC marks throughout the entire genome. Alternatively, 5hmC can be indirectly identified using an oxidative bisulfite sequencing (OXBS-seq) approach where 5hmC is oxidized into 5-formyl-cytosine (5fC), which is then converted into T during bisulfite treatment (Booth et al., 2012). In other words, only 5mC remains unconverted (C), while 5hmC and C become converted into T (Fig 4.1). This is referred to as an *indirect* profiling of 5hmC, because it relies on a comparison with a parallel BS-seq where both 5mC and 5hmC are unconverted in order to accurately quantify 5hmC levels at single nucleotide resolution. By subtracting methylation levels of OXBS-seq from BS-seq, one can infer the starting level of 5hmC within the sample. Despite drawbacks, including the requirement of parallel and costly sequencing, OXBS is still the only available method that offers genome-wide, simultaneous profiling of both C, 5mC and 5hmC without relying on enzymatic activity or affinity of any recombinant proteins.

Since the recent invention of OXBS, the technology has been utilized in various mammalian contexts. In humans, during monocyte to macrophage differentiation, OXBS detected transient 5hmC levels during 5mC demethylation at loci specifically associated with genes unique to macrophages. These regions later become nucleosome-free and actively transcribed in macrophages (Wallner et al., 2016). Similarly, during spermatogonia stem cell differentiation, OXBS revealed partially and differentially methylated domains associated with spermatogenesis (Kubo et al., 2015). In a mouse model for aging, OXBS enabled the detection of small but significant differences in 5hmC levels in age-related genes on the X-chromosome of young and old hippocampi (Hadad et al., 2016). In cancer models, genome-wide BS/OXBS profiling revealed global hypermethylation and decreased 5hmC levels in human glioblastoma when compared to normal tissue (Raiber et al., 2017).

While OXBS technology has been applied to 5hmC detection in several vertebrate systems, combined methylome and hydroxymethylome profiling in the retina has not been performed in any model organism. This is important because genes involved in retinal development such as specification TFs and genes encoding cell intrinsic and extrinsic signaling molecules are likely epigenetically regulated at the level of DNA methylation and hydroxymethylation. Changes in 5mC and/or 5hmC levels at or near these genes may correlate with changes in gene expression, and as shown in Chapter 3, misregulation of hydroxymethylation can lead to misexpression of critical genes, resulting in failure of cellular differentiation.

To fully understand the role DNA hydroxymethylation plays a role in the epigenetic control of retinal development, it is necessary to generate comprehensive 5mC/5hmC profiles of individual retinal cell types at relevant stages during retinogenesis. Here we present the first genome-wide map of DNA methylation (methylome) and

hydroxymethylation (hydroxymethylome) at single nucleotide resolution in the developing embryonic retina. Using zebrafish as a model, we isolated a pure population of retinal progenitors (RPCs) at specific time points before and during differentiation. Our analysis identifies numerous regions that display differential methylation and hydroxymethylation levels, many of which are overlapping. Some of these are associated with known eye genes while others are associated with potentially novel genes involved in retinal development.

4.2 RESULTS

4.2.1 Isolation of genomic DNA from a pure RPC population

To obtain a pure population of RPCs, we utilized a transgenic reporter zebrafish carrying *vsx2:*GFP, which expresses a strong GFP signal starting at 20hpf in RPCs (Vitorino et al., 2009). RPCs were collected via FACS from eyes dissected at two time points: early (22hpf), when most of the retina is composed of actively proliferating and undifferentiated RPCs; and late (27hpf), immediately before the onset of the first differentiation event, when RPCs exit the cell cycle and become RGCs (Hu and Easter, 1999; Schmitt and Dowling, 1996). To ensure a viable and correctly identified population, cells were sorted based on size, GFP+ signal, and the absence of cell death signal (Live/Dead staining) (Fig 4.1A). To independently verify the collection of the appropriate cell population, sorted cells were visually inspected under a fluorescent microscope and a fraction of the cells were processed for semi-quantitative RT-PCR. RPC populations were found to retain expression of GFP and *vsx2* for at least 2 hours post-sorting, while GFP- non-RPC populations, likely extra-ocular tissue, did not possess GFP or *vsx2* transcripts at a detectable level (data not shown). In this pilot study, we also

collected cells from whole embryos as a reference sample. Altogether, two comparisons were done: 22hpf RPC *vs* whole embryo, and 22hpf RPC *vs* 27hpf RPC.

Genomic DNA was extracted, quantified, and processed through OXBS and BS reactions in parallel (Fig 4.1B and Appendix A). Before processing, genomic DNA was spiked with two sets of quality control spike-in DNA duplexes containing 5mC and 5hmC at known locations: digestion control (DC) for interrogation of the OXBS reaction before sequencing, and sequencing control (SQ) for post-sequencing assessment of conversion levels. For the OXBS reaction, DNA was treated with an oxidizing agent followed by a bisulfite conversion treatment. For the BS reaction, DNA was mock-treated with water then bisulfite converted. Treated DNA was sub-sampled to assess the completion of the OXBS reaction by amplification of the DC spike-in followed by digestion with the restriction enzyme TaqaI, which cuts at the TCGA sequence (also present in the cutting control: CC). The DC duplex contains T^{5hm}CGA which is fully converted into TTGA (non-digestible) after OXBS treatment, while in BS treatment it remains as TCGA and will thus be fully digested (Fig 4.2A). This amplification-digestion assay provided a semi-quantitative assessment of the BS/OXBS reaction efficiency pre-sequencing and allowed us to decide whether to proceed with next-generation sequencing based on conversion efficiency. Because we observed a distinct size difference between fully digested and undigested fragments post-treatment (Fig 4.2A, arrows), the libraries were thus processed for sequencing.

4.2.2 Methylome and Hydroxymethylome profiling approach

Once the samples passed the DC amplification-digestion checkpoint, genomic DNA samples post-BS and post-OXBS treatment were processed through adapter ligation and indexing, then sequenced on an Illumina NextSeq using paired-end 150 cycles. Using FastQC for an initial quality check, we observed a good sequence quality throughout the entire length of the reads. In addition, overall sequences exhibited a low level (<2%) of C/G nucleotides and a high level of A/T nucleotides, typical of a bisulfite converted libraries (Fig 4.2B-C).

To quantitatively determine the level of base conversion after BS and OXBS treatment, sequences were aligned to the SQ spike-in control duplexes and the levels of C, 5mC, and 5hmC were quantified. Overall, we observed a complete conversion ratio, where all unmethylated Cs were converted to Ts, and 5mCs + 5hmCs remained unconverted after BS treatment. In contrast, after OXBS treatment, only 5mCs remained unconverted, and all Cs+5hmCs were all converted to Ts (Fig 4.2D-E).

Raw sequences (fastq.gz), excluding the spike-in sequences, were aligned to the fully converted and indexed zebrafish genome (GRCz10) using a bisulfite compatible aligner (Pedersen et al., 2014). After alignment, the individual methylation score at each cytosine base was extracted using Bismark Methylation Extractor (Krueger and Andrews, 2011) at a mean depth of 1.9 fold coverage per cytosine. It is important to note that the methylation call files generated here contain combined 5mC+5hmC scores for BS-seq and only 5mC scores for OXBS-seq. To generate an estimated level of individual C, 5mC, and 5hmC at every nucleotide position, we relied on a maximum likelihood estimation approach (Qu et al., 2013), utilizing MLML (Maximum-Likelihood

Methylation Level) software, which takes into account methylation counts from both BS/OXBS and the depth of read coverage; MLML-based estimation improves in accuracy as the read coverage increases. Once the estimation was performed for each sample, three methylation call files were generated. Each C was assigned a percentage (0-100%) for each possible modification: 5mC, 5hmC, and unmethylated C. Overall, we observed an average of 82.79% combined methylation level across all CpGs in the genome. Of these, 74.98% are 5mC and 7.81% are 5hmC.

From this step forward, the individual 5mC/5hmC percentage files can be treated as a regular methylation score file and are thus compatible with a wide range of downstream methylation analysis pipelines. Here we use the MethPipe software suite (Song et al., 2014) to identify regions of hypo-methylation and hyper-methylation (HMRs), differentially methylated regions (DMRs), and differentially hydroxymethylated regions (DhMRs). Identified regions were exported as scored genomic coordinates (.bed files) similar to an output from a ChIP-seq experiment, then annotated and motif analyzed using Homer (Heinz et al., 2010).

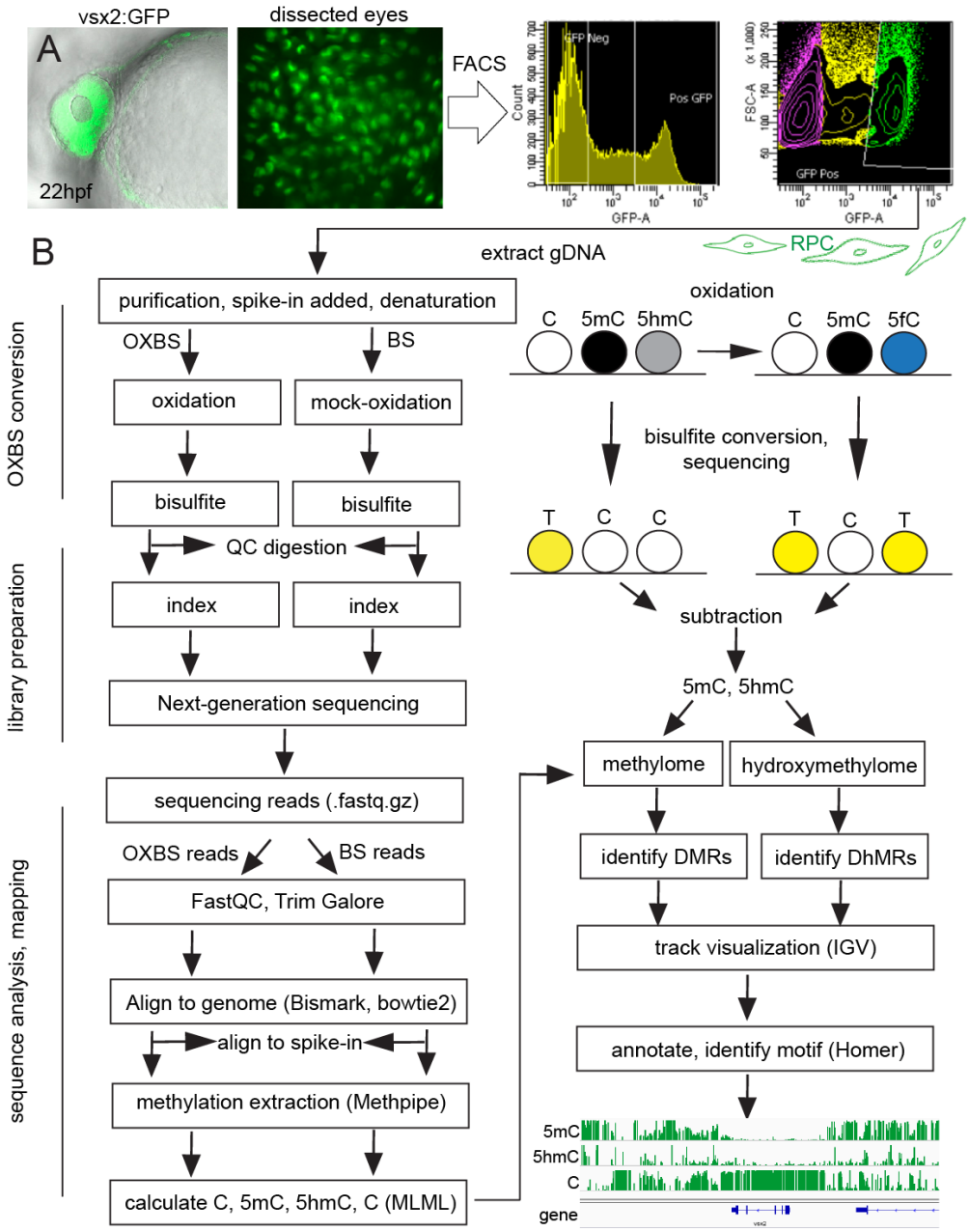

Figure 4.1: Workflow of methylome and hydroxymethylome profiling in the retina. Zebrafish vsx2:GFP embryos at 22hpf were dissected to obtain a minimum of 200 eyes, which were dissociated and cell sorted for GFP+ RPC population (A). Genomic DNA was extracted, divided in half, and processed through a parallel BS/OXBS library preparation protocol, where 5hmC is either remaining as C in the BS half or converted to 5fC and then T in the OXBS half. All reads were mapped, 5mC/5hmC levels calculated, and genome-wide maps generated (B).

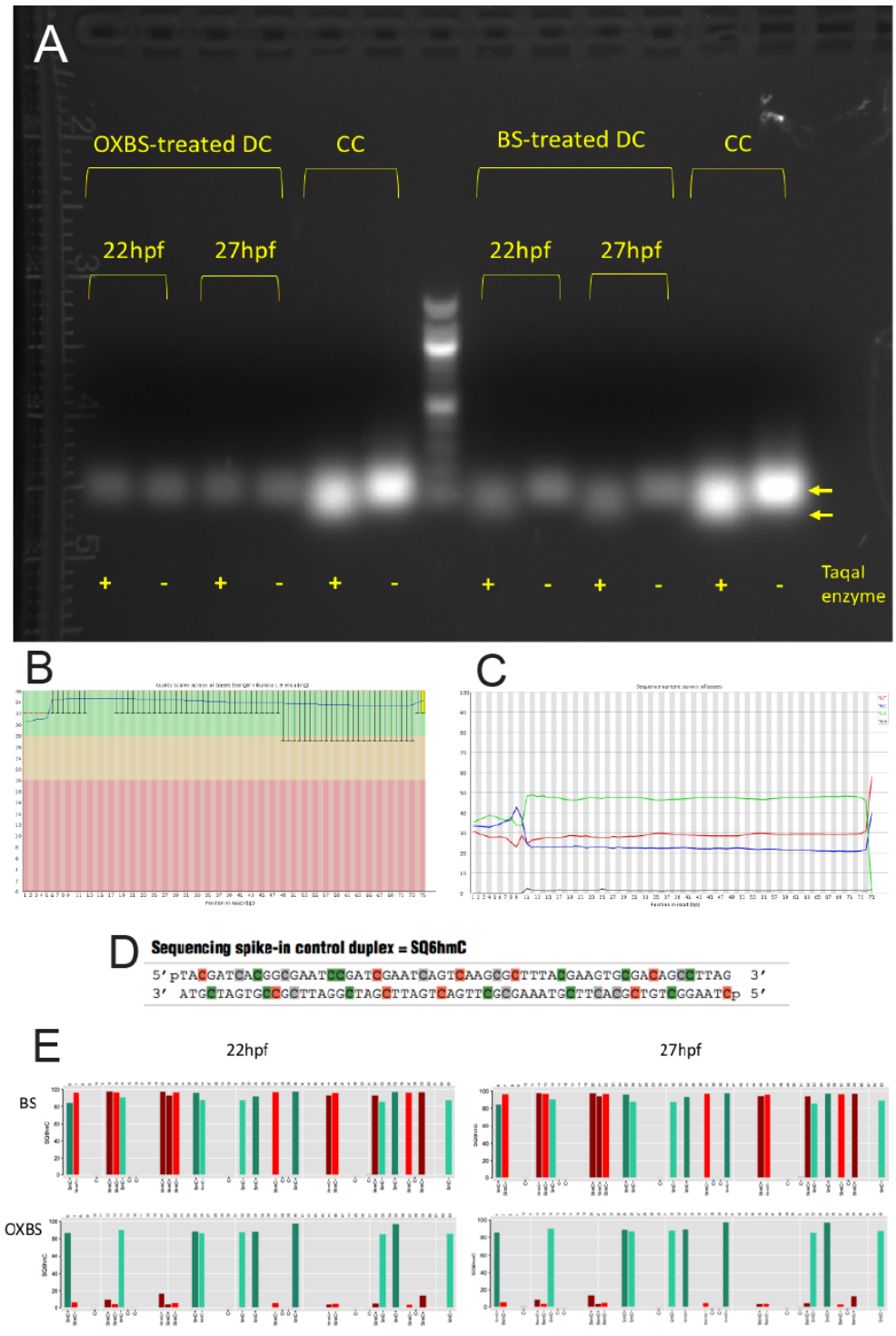

Figure 4.2: Digestion control (DC) and sequencing control (SQ) interrogation show complete C and 5hmC conversion after OXBS treatment.

(**Figure 4.2 – continued**). Genomic DNA samples were spiked with DC and SQ duplexes before BS/OXBS parallel reactions and library preparation. Amplification-digestion followed by gel electrophoresis semi-quantitatively show complete digestion by TaqaI after OXBS treatment, indicating that 5hmC is fully converted to T (A). CC: cutting control. SQ6hmC: sequencing control containing six 5hmC nucleotides. Sequences generated by Illumina NextSeq revealed high quality reads (B) and low GC content (C). Post-sequencing alignment quantitatively showed a complete C and 5hmC conversion after OXBS and only C→T conversion after BS reaction. (D) Sequencing control SQ6hmC spike-in duplex. (E) Quantitative measurement of 5mC/5hmC levels on SQ6hmC post-sequencing. Red bars: 5hmC; Green bar: 5hmC.

4.2.3 Early RPCs (22hpf) to whole embryo comparison

We sought to perform the first methylome and hydroxymethylome comparison between 22hpf RPCs and whole embryos 1) because these cell/tissue types are among the most abundant and 2) to test whether we can detect changes in 5mC and 5hmC between groups of cells that are biologically vastly different. Using the extracted methylation scores, we observed a high average genome-wide methylation level of 82% (5mC+5hmC) which is consistent with studies using other vertebrate model systems (e.g. (Mo et al., 2016; Okae et al., 2014; Ziller et al., 2013)).

To spot-check the data, individual 5mC, 5hmC, and unmethylated C tracks were visualized at candidate loci where methylation statuses are already known using Integrative Genome Viewer (IGV, Broad Institute). These tracks were then compared to published WGBS data (which does not distinguish between 5mC and 5hmC) from embryonic mouse and human whole retinal tissues (Pecan database (Aldiri et al., 2017), https://pecan.stjude.cloud/proteinpaint/study/retina2017). We observed a large (80kb) domain covering nine *hox* genes where 5mC and 5hmC are nearly absent in both 22hpf RPCs and whole embryos (Fig 4.3A). Notably, the *hox* gene cluster also possesses a similarly large and hypomethylated domain in both mouse and human embryonic retinae.

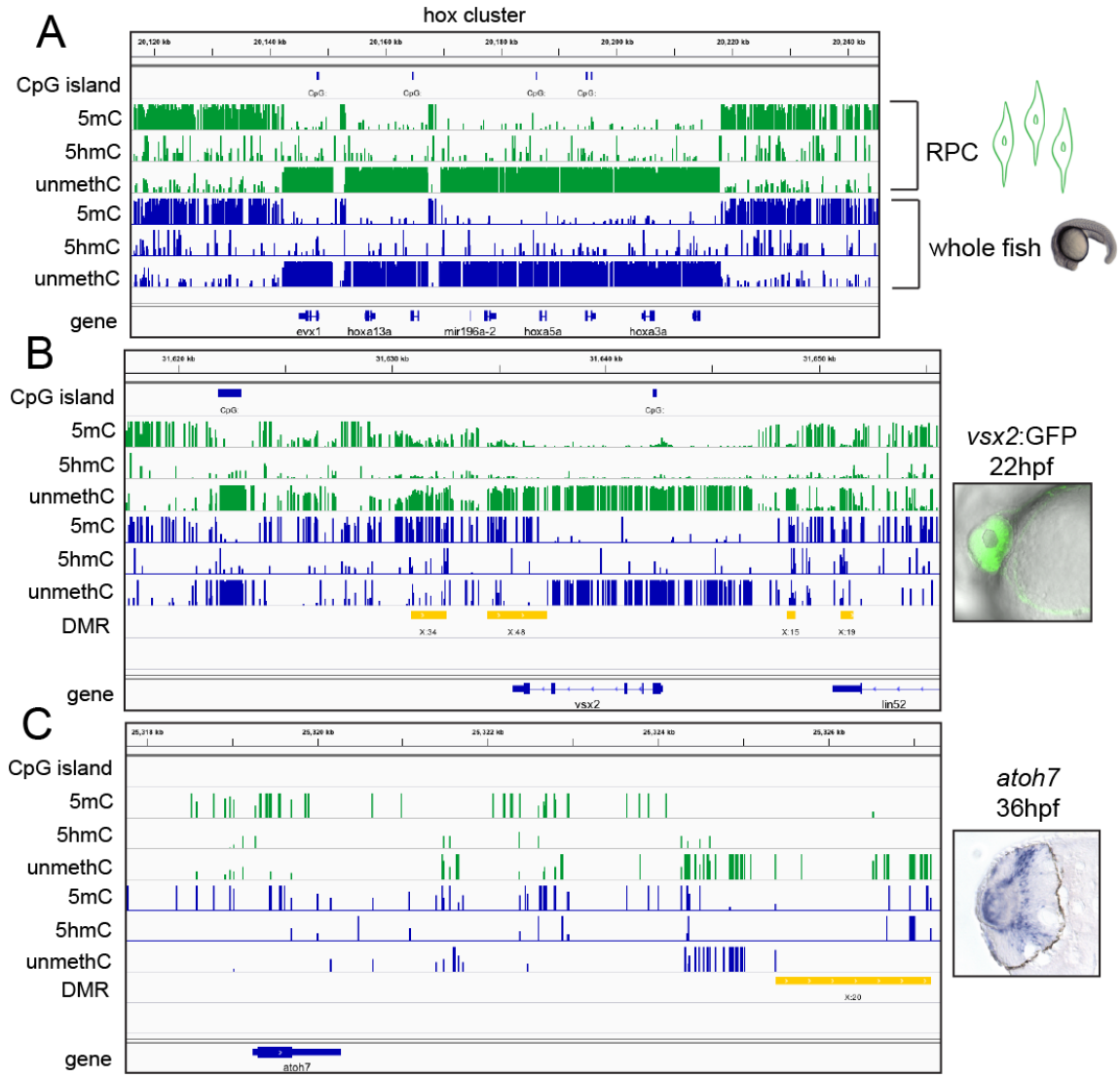

Figure 4.3: Methylome and hydroxymethylome profiles show expected 5mC and 5hmC distribution patterns at known genomic regions. Windows encompassing *vsx2*, *atoh7*, and hox cluster were captured using IGV. Methylation scores are displayed as a percentage (0-100%) for each cytosine modification (5mC, 5hmC, unmethylated C) in each track. Green: 22hpf RPC; Blue: whole embryo; Yellow boxes: differentially methylated regions (DMRs).

To locate regions where 5mC and 5hmC levels significantly vary between 22hpf RPCs and whole embryos, 56,524 DMRs and 4,958 DhMRs were identified using MethDiff (Song et al., 2014). Of these, 4,285 DMRs and 724 DhMRs possess at least 5 CpGs and 3 significantly differentially methylated CpGs ($p<0.05$). To test the biological relevance of these regions, we spot-checked these data using loci at/near retinal developmental genes. First we analyzed *vsx2*, a gene highly expressed in RPCs, and found a 5mC methylation valley approximately 10kb upstream and downstream of the transcription start site (TSS), a feature commonly observed in actively transcribed genes in human and mouse (Aldiri et al., 2017). This hypomethylated *vsx2* region is present in both 22hpf RPCs and whole embryos, likely because *vsx2* is also expressed at a lower level in the brain and spinal cord, in addition to the strong expression in RPCs (Vitorino et al., 2009). Interestingly, there are four DMRs surrounding the *vsx2* TSS where 5mC levels are lower in 22hpf RPCs than whole embryos, correlating with higher *vsx2* expression in RPCs (Fig 4.3B). Second, *atoh7*, a specification factor for RGCs, the earliest-born retinal cell type, shows 1) a significant DMR within 5kb of the gene where 5mC levels are lower in 22hpf RPCs than in whole embryos, and 2) a few sparse positions where 5hmC levels are higher in 22hpf RPCs than in the whole embryos (Fig 4.3C). Collectively, these positions of lower 5mC and higher 5hmC may indicate pre-activation of *atoh7* prior to active transcription later in development when RPCs begin to differentiate into RGCs starting at 28hpf (Hu and Easter, 1999).

Next, to take an unbiased approach and identify novel genes that may be involved in retinal development, 204 DMR/DhMR overlapping regions were identified where 5mC is lower and 5hmC is higher in RPCs than in the whole embryos (Fig 4.4A and Table 4.1). We reasoned that regions where 5mC is lower and 5hmC is higher in RPCs than in whole embryos are likely located near genes that are or about to be more actively

transcribed in RPCs than in the embryo as a whole. Interestingly, many of these DMR/DhMR intersecting regions ("peaks") are located within transposable elements (TE) and long terminal repeat (LTR) regions. To specifically test whether these peaks are indeed enriched for TE and LTR, we 'shuffled' the genomic coordinates of the peaks so that they are randomly distributed throughout the genome and compared their enrichment on known genomic features. Indeed, DMR/DhMR intersecting peaks show higher enrichment with LTR and TE than the shuffled regions (Fig 4.4B-C), indicating that these elements may be epigenetically regulated, a known phenomenon that has been extensively described in the context of DNA methylation in other model organisms (Daron and Slotkin, 2017; Slotkin and Martienssen, 2007).

Many of these 204 DMR/DhMR intersecting regions are located near known eye genes while some are near unknown or uncharacterized genes (Table 4.1). For example, the top candidate is located upstream of *microRNA-181a* (Fig 4.4D), which is one of the most biologically interesting candidates because *mir181a* is not expressed in RPCs at 22hpf as shown by our RNA-seq data (Table 4.1); however, it is highly and specifically expressed later at 4dpf in zebrafish RGCs (Wienholds, 2005) (Fig 4.4E). In mouse, *miR181a* has been shown to promote de-differentiation of fibroblast to pluripotent stem cell, in conjunction with other pluripotency factors (Judson et al., 2013). This suggests that *miR181a* may play a role in cellular differentiation during development, but its role has not yet been functionally tested in the retina. Similar to *miR181a*, other genes near DMR/DhMR intersecting regions are also expressed at low levels (Table 4.1), indicating that they may be undergoing pre-activation, although expression profiling at a later time point and in differentiated cells is necessary to support this hypothesis.

To identify potential roles of the 204 DMR/DhMR intersecting regions, we performed motif analysis using Homer and gene ontology analysis using DAVID, which

revealed that genes located near DMR/DhMR peaks are likely transcription factors (TFs), with potential roles in development and cell differentiation. Additionally, the top candidate motifs correspond to binding sites for TFs with well-established roles in eye development, such as Pou4f3 and Six6 (Fig 4.4F-G).

It is important to note that these analyses were generated from one BS/OXBS sequencing run, and additional replicates are required in order for these results to be more reliable and conclusive. Additional rounds of sequencing at higher depth, although costly to run, will allow statistical analyses such as false discovery rate (FDR) that cannot be calculated from a single data set.

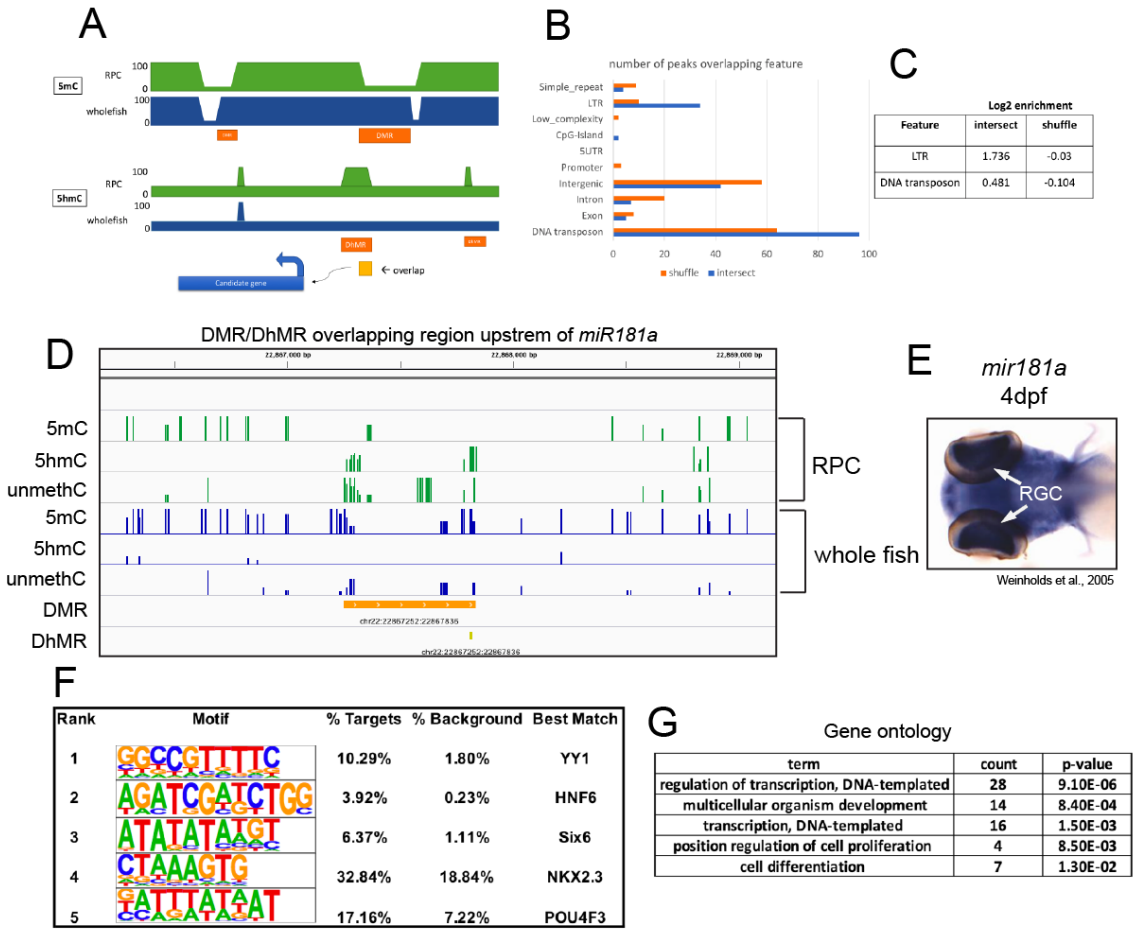

Figure 4.4: DMR/DhMR intersecting regions are overlapping transposons and located near novel retinal genes. DMR/DhMR intersecting regions (peaks) were located and nearest genes identified (A). Intersecting peaks show higher enrichment score for LTR and transposons, compared to randomly distributed peaks (B-C). The top candidate peak is located upstream of *miR181a*, which is expressed specifically in RGCs (D-E). GO and Motif analyses show that intersecting peaks are enriched for transcription factors involved in development (F-G).

Gene Name	Gene Description	Normalized score	Expression domain	RNA-seq (TPM)
mir181a-1	microRNA 181a-1	0.50	eye, GC, BP	0.00
fars2	phenylalanyl-tRNA synthetase 2, mitochondrial	0.38	RPE	4.37
zgc:174354	zgc:174354	0.32	no data	-
si:rp71-1h3.1	si:rp71-1h3.1	0.29	no data	-
znf1070	zinc finger protein 1070	0.28	no data	3.67
nppa	natriuretic peptide A	0.25	heart	0.32
cryba1l1	crystallin, beta A1, like 1	0.23	lens	8.70
p2ry10	purinergic receptor P2Y, G-protein coupled, 10	0.22	no data	0.00
zgc:172282	zgc:172282	0.21	no data	0.00
cfap58	cilia and flagella associated protein 58	0.20	no data	0.54
tbxa2r	thromboxane A2 receptor	0.19	eye, brain	0.00
igfbp2b	insulin-like growth factor binding protein 2b	0.19	eye	2.42
si:dkey-27n6.1	si:dkey-27n6.1	0.19	no data	0.32
fgfrl1b	fibroblast growth factor receptor-like 1b	0.18	eye, lens	2.88
prkag3b	protein kinase, AMP-activated, gamma 3b non-catalytic subunit	0.18	no data	0.38
otx1b	orthodenticle homeobox 1b	0.17	retina	6.19
sgut1	SGT1 homolog, MIS12 kinetochore complex assembly cochaperone	0.17	eye, brain	14.89
xylb	xylulokinase homolog (H. influenzae)	0.17	liver	7.59
cnot6b	CCR4-NOT transcription complex, subunit 6b	0.17	no data	5.26

Table 4.1: List of the top 20 genes containing nearby differentially methylated regions (DMRs) overlapping with differentially hydroxymethylated regions (DhMRs) between 22hpf RPCs and whole embryos, ranked by normalized score (number of significantly differentially methylated CpGs divided by the region size). Expression domains are summarized from the Zebrafish Information Network (zfin.org) either from direct submissions or curated publications. Gene expression levels were calculated from 22hpf RPC RNA-seq and reported in transcripts per kilobase million (TPM).

4.2.4 Comparison between early RPCs (22hpf) and late RPCs (27hpf)

To further investigate methylome and hydroxymethylome dynamics at a pivotal time window of retinal development, we performed BS/OXBS-seq in early RPCs (22hpf) and compared 5mC/5hmC levels to late RPCs (27hpf), immediately before the onset of neurogenesis events when the first RGCs exit the cell cycle and begin to differentiate (Hu and Easter, 1999). Genomic DNA from FAC-sorted cells were BS/OXBS-treated, libraries generated and sequenced at an average coverage depth of 2.12 fold. Initial analysis based on gene-by-gene observation of 5mC/5hmC profiles at known retinal genes and *hox* clusters showed no noticeable differences between the profiles at any of the candidate regions. Entire *hox* clusters were found to be largely hypomethylated in both 22hpf and 27hpf RPCs (Fig 4.5A). Similarly, 5mC/5hmC profiles appear identical at candidate retinal neurogenesis genes in both 22hpf and 27hpf RPCs: *vsx2* and *ptf1a* both show hypomethylated 5mC regions throughout the entire gene body, while *crx* is hypomethylated at the TSS, and *atoh7* remains methylated (Fig 4.5B). Importantly, no significant DMRs or DhMRs were present in or near these five genomic windows, indicating that the methylome and hydroxymethylome profiles of these regions are virtually identical.

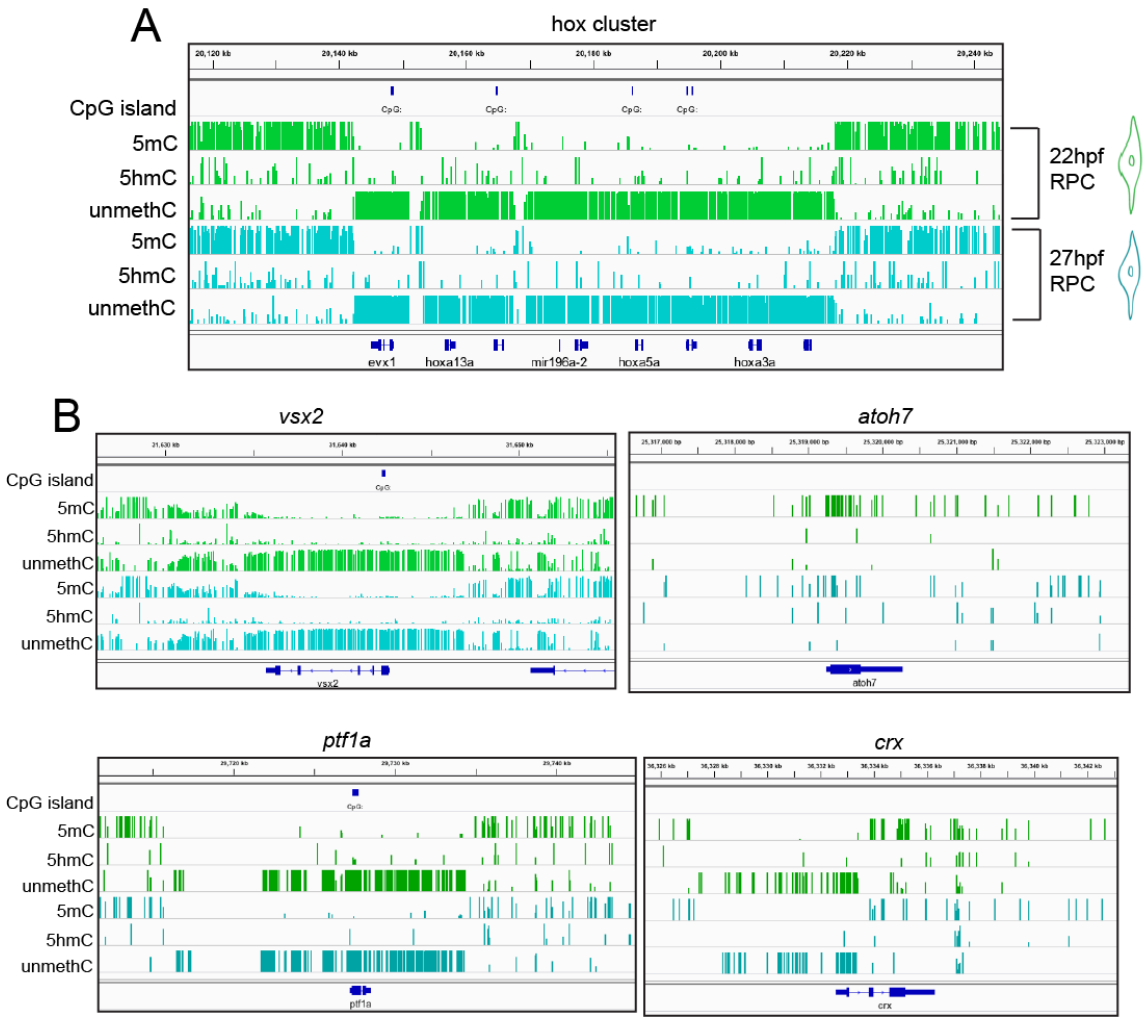

Figure 4.5: Methylome profiles appear similar between hox cluster and candidate retinal neurogenesis genes of 22hpf and 27hpf RPC. Methylation scores are displayed in percentage (0-100%) for each cytosine modification (5mC, 5hmC, unmethylated C) in each track. Green: 22hpf RPC; Cyan: 27hpf RPC.

Despite having overall similar epigenomic profiles at candidate retinal genes, differential methylation analysis identified 3,836 DhMRs and 53,938 DMRs between 22hpf and 27hpf RPC populations throughout the genome. Of these, 868 DhMRs and 1,521 DMRs possess at least 5 CpGs and 3 significantly differentially methylated CpGs ($p<0.05$). These regions are categorized, based on the changes in 5mC and 5hmC levels between 22hpf and 27hpf, into four categories: 5mC-losing (n=1,100), 5mC-gaining (n=421), 5hmC-gaining (n=449), and 5hmC-losing (n=419). To determine the biological relevance of these regions, we performed annotation based on the location of closest genes and performed motif and GO analyses (Fig 4.6A-C).

We first investigated genes located near 5hmC-gaining regions because these are likely to become actively transcribed at or soon after 27hpf (Table 4.2). Genes in this group are composed of developmental transcription factors and signaling molecules with known roles in development (e.g. *pou3f1, rbp2a, crabp2a, smad7*). Closer inspection of the methylation score tracks showed that these regions indeed gained 5hmC levels in RPCs over time from 22hpf to 27hpf (Fig 4.6A). Interestingly, based on the publicly available expression data (zfin.org), five of the top 20 genes are expressed within the eye between 32-42hpf, and one, *rbp2a*, is expressed specifically within differentiated neurons, presumably early-born photoreceptor cells in the ventral retina later in development (Fig 4.7A).

To identify genes undergoing activation and becoming transcribed, which often contain regions of low 5mC and elevated 5hmC as shown in the case of *miR181* above, we intersected the 5mC-losing regions with 5hmC-gaining regions, and again performed annotation, motif analysis, and GO analysis (Fig 4.6D-F). While there are overlapping

genes between this list and the 5hmC-gaining list (e.g. *smad7, crabp1a, pth1b*), many of the identified genes possess unique characteristics (Table 4.3). First, their closest DMR/DhMR peaks are enriched for vsx2 binding motifs, suggesting that they may be regulated by vsx2 (Fig 4.6E). Second, GO analysis showed enrichment for eye, retina, and optic nerve developmental terms, which implies that genes containing these peaks may be involved in RGC differentiation (Fig 4.6F). Third, *in situ* expression of at least 3 of the top 20 genes (*slc17a6b, glra4b, pcbp4*) is detected in the eye, with *slc17a6b* being expressed specifically in RGCs (Fig 4.7B).

Next, we hypothesized that genes that are no longer required once RPCs differentiate into neurons would undergo inactivation and begin to gain 5mC over the course of 22-27hpf. These genes may be expressed in non-RPC cells/tissues later in development, but their expression should be absent in RPC-derived retinal cells. To identify these genes, we annotated and performed GO and motif analyses for DMRs where 5mC levels are higher in 27hpf than 22hpf RPC (5mC-gaining) (Fig 4.6G-I). Indeed, several genes in this category are expressed in non-ocular tissues, such as head mesoderm, somites, cranial muscle, cartilage, and olfactory receptors at post-27hpf time points (Table 4.4). Interestingly, GO terms for genes in this group contain a diverse mix of biological processes such as homophilic cell adhesion, chromatin modification, and epithelial cell apical/basal polarity, which may coincide with the loss of RPC identity as a pseudostratified epithelium to become differentiated neurons after 27hpf.

It is important to note that because the expression data of the remaining genes on these lists are not yet available on the public database or are available at resolutions/regions that are uninformative, we cannot conclude that they are not

expressed in the cells/tissues of interest. Additional factors should be taken into consideration when interpreting this dataset: 1) these results are still preliminary and additional sequencing at higher depth in biological replicates should improve the data quality 2) DMRs and DhMRs will be intersected with lists of differentially expressed genes in 27hpf RPCs from RNAseq, which will be performed in the near future 3) candidate DMRs and DhMRs can be independently validated by traditional gene-by-gene bisulfite sequencing for combined 5mC+5hmC and glucosylation-digestion (Quest assay performed in Chapter 3) for 5hmC specifically.

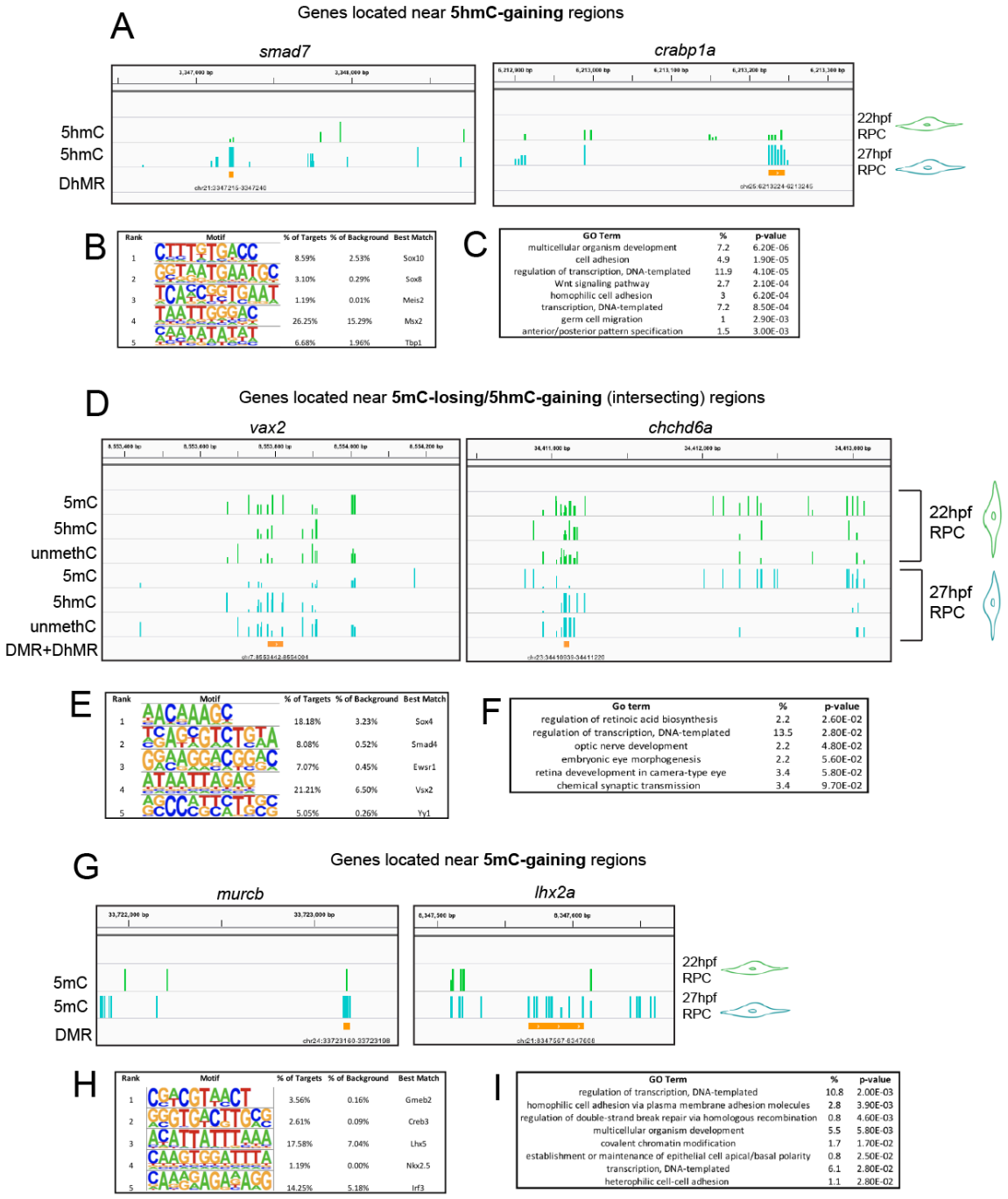

Figure 4.6: Regions of differential methylation and hydroxymethylation are located near genes with known and unknown roles during retinal development and are potential binding sites for retinal developmental transcription factors.

Gene Name	Gene Description	Normalized score	Expression
rbp2a	retinol binding protein 2a, cellular	0.44	PR, ventral patch
hdac3	histone deacetylase 3	0.43	head, ubiquitous
sdc4	syndecan 4	0.38	NCC, melanocyte
htra3a	HtrA serine peptidase 3a	0.36	head, ubiquitous
nanos3	nanos homolog 3	0.31	PGCs
pth1b	parathyroid hormone 1b	0.31	eye, forebrain, heart
znf236	zinc finger protein 236	0.30	ubiquitous
crabp1a	cellular retinoic acid binding protein 1a	0.30	GC, PR
scn12aa	sodium channel, voltage gated, type XII, alpha a	0.30	eye, brain spinal cord
txnl1	thioredoxin-like 1	0.29	ubiquitous
smad7	SMAD family member 7	0.29	eye, RPC
pou3f1	POU class 3 homeobox 1	0.29	eye, brain
plcxd3	phosphatidylinositol-specific phospholipase C, X domain containing 3	0.28	brain, branchial arch
park2	parkin RBR E3 ubiquitin protein ligase	0.25	ubiquitous
otud6b	OTU domain containing 6B	0.25	ubiquitous
si:ch211-262n1.4	si:ch211-262n1.4	0.25	no data
grm1b	glutamate receptor, metabotropic 1b	0.24	RGC, optic tectum
zgc:103482	zgc:103482	0.23	no data
zgc:174710	zgc:174710	0.23	no data
adck1	aarF domain containing kinase 1	0.22	no data

Table 4.2: List of the top 20 genes containing a nearby regions where 5hmC levels are significantly higher in 27hpf than 22hpf RPCs (5hmC-gaining) ranked by normalized score (number of significantly differentially hydroxymethylated CpGs divided by the region size). Expression domains are summarized from the Zebrafish Information Network (zfin.org) either from direct submissions or curated publications.

Gene Name	Gene Description	Normalized score	Expression
crabp1a	cellular retinoic acid binding protein 1a	0.45	eye, optic tectum
sdc4	syndecan 4	0.38	crest, melanocyte
pth1b	parathyroid hormone 1b	0.31	eye, forebrain
nol4lb	nucleolar protein 4-like b	0.26	ubiquitous
slc17a6b	solute carrier family 17 (vesicular glutamate transporter), member 6b	0.25	RGC, brain
smad7	SMAD family member 7	0.25	eye (RPC?), brain
egln3	egl-9 family hypoxia-inducible factor 3	0.24	ubiquitous
zgc:174710	zgc:174710	0.23	no data
eif3ja	eukaryotic translation initiation factor 3, subunit Ja	0.21	eye and tectum
chchd6a	coiled-coil-helix-coiled-coil-helix domain containing 6a	0.19	eye and brain
tubb2	tubulin, beta 2A class IIa	0.19	eye, CMZ, brain
si:ch211-262n1.4	si:ch211-262n1.4	0.19	no data
gbx1	gastrulation brain homeobox 1	0.16	mibrain, hindbrain
prkrip1	PRKR interacting protein 1 (IL11 inducible)	0.14	no data
tmem129	transmembrane protein 129, E3 ubiquitin protein ligase	0.13	eye, brain, pectoral fin
vax2	ventral anterior homeobox 2	0.13	ventral retina
gsdf	gonadal somatic cell derived factor	0.13	testis, ovray
glra4b	glycine receptor, alpha 4b	0.12	retina (RGC, INL)
wbp1	WW domain binding protein 1	0.11	ubiquitous
pcbp4	poly(rC) binding protein 4	0.11	cranial ganglion

Table 4.3: List of the top 20 genes containing nearby differentially methylated regions (DMRs) overlapping with differentially hydroxymethylated regions (DhMRs) between 22hpf and 27hpf RPCs, ranked by normalized score (number of significantly differentially methylated CpGs divided by the region size). Expression domains are summarized from the Zebrafish Information Network (zfin.org) either from direct submissions or curated publications.

Gene Name	Gene Description	Normalized score	Expression
si:dkey-61p9.11	si:dkey-61p9.11	0.30	no data
lhx2a	LIM homeobox 2a	0.23	neural tube, olfactory organ
murcb	muscle-related coiled-coil protein b	0.22	muscle, bone
lgi3	leucine-rich repeat LGI family, member 3	0.21	no data
znf1068	zinc finger protein 1068	0.20	no data
ntn1a	netrin 1a	0.18	CNS, optic stalk
colec11	collectin sub-family member 11	0.18	head mesoderm, liver
rpl27a	ribosomal protein L27a	0.17	ubiquitous
them4	thioesterase superfamily member 4	0.15	no data
stoml2	stomatin (EPB72)-like 2	0.14	no data
ints2	integrator complex subunit 2	0.14	ubiquitous
si:dkey-169i5.4	si:dkey-169i5.4	0.14	no data
hrasb	Harvey rat sarcoma viral oncogene homolog b	0.13	no data
ing4	inhibitor of growth family, member 4	0.12	ubiquitous
fgf7	fibroblast growth factor 7	0.11	no data
slc39a4	solute carrier family 39 (zinc transporter), member 4	0.11	ubiquitous
LOC100149406	odorant receptor	0.10	olfactory organ
tox2	TOX high mobility group box family member 2	0.10	no data
crb2b	crumbs family member 2b	0.10	rod, cone, pineal gland
znf609a	zinc finger protein 609a	0.10	no data

Table 4.4: List of the top 20 genes containing nearby regions where 5mC levels are significantly higher in 27hpf than 22hpf RPCs (5mC-gaining), ranked by normalized score (number of significantly differentially methylated CpGs divided by the region size). Expression domains are summarized from the Zebrafish Information Network (zfin.org) either from direct submissions or curated publications.

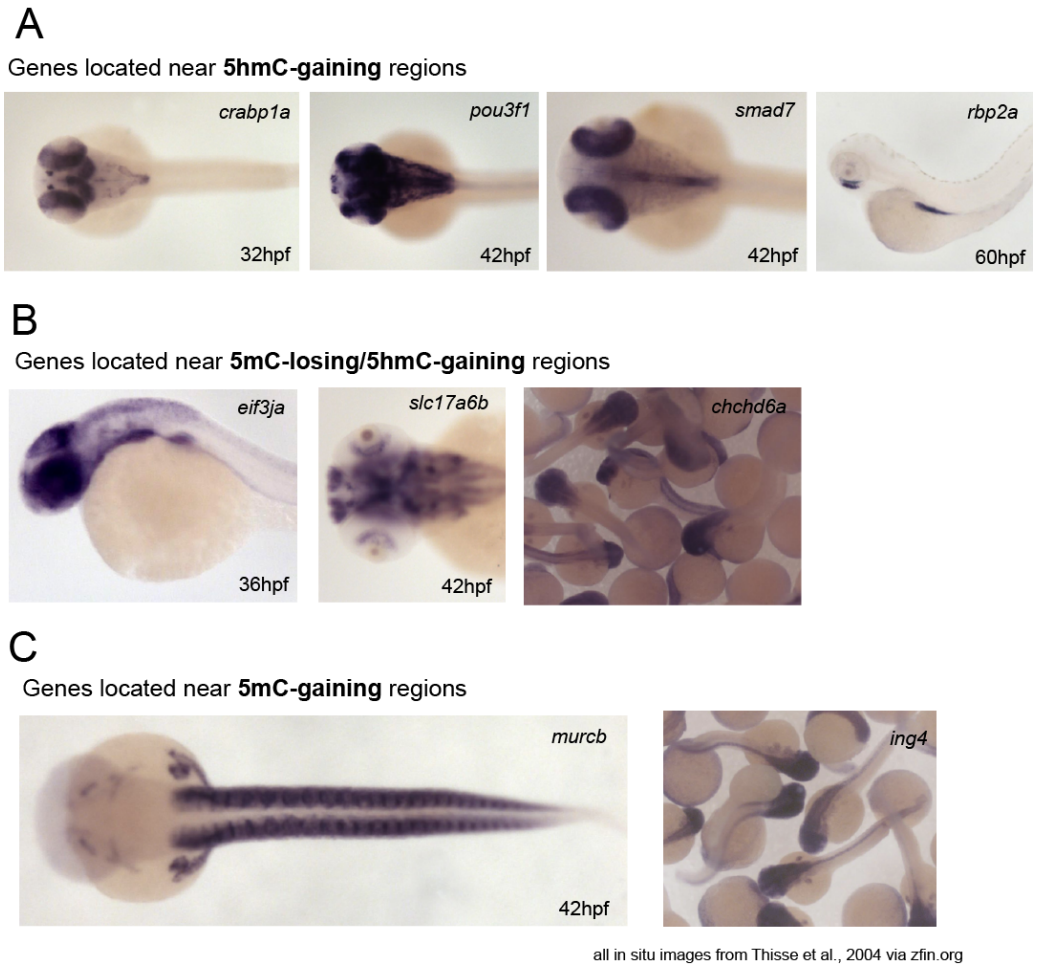

Figure 4.7: Methylome and hydroxymethylome analysis of 22hpf vs 27hpf RPCs reveals candidate retinal developmental genes with expression in relevant domains. For 5hmC-gaining genes, *crabp1a*, *pou3f1*, and *smad7* are expressed in the RPCs and brain, while *rbp2a* is expressed in the PR ventral patch (A). For 5mC-losing/5hmC-gaining genes, *eif3ja* is expressed in the eye and brain; *slc17a6b* is expressed in the RGCs, and *chchd6a* is ubiquitously expressed (B). For 5mC-gaining genes, *murcb* is expressed in the muscle and cartilage, and *ing4* is expressed ubiquitously (C). All images are from the publicly available database, Zebrafish Information Network (zfin.org) (Thisse and Thisse, 2004).

4.3 DISCUSSION

Simultaneous profiling of methylation and hydroxymethylation marks in a pure cell population has long been a challenge in developmental epigenetics. Traditional WGBS-seq offers a massive amount of mixed methylation data points where 5mC and 5hmC cannot be distinguished, while affinity-based sequencing offers specific 5mC/5hmC in a regional but not a base-resolution view of the epigenome. Capitalizing on the recent development of BS/OXBS chemistry, coupled with next-generation DNA sequencing, we show here for the first time a complete methylome/hydroxymethylome map of isolated RPCs during retinal development at nucleotide-resolution. While our overarching goal is to profile 5mC/5hmC of RPCs and all differentiated retinal cell types over the course of development, we present here the data generated from early RPCs, late RPCs, and whole embryos, representing the first phase of this endeavor.

By comparing the methylome and hydroxymethylome of early (22hpf) RPCs to whole embryos at the matched time point, we verified that the 5mC/5hmC profiles of candidate retinal genes (e.g. *vsx2*, *atoh7*) follow the expected trajectories according to their known expression patterns and timing in RPCs and differentiated neurons. Using an unbiased approach, we identified several regions of differential methylation and hydroxymethylation (DMR, DhMRs) that are located near genes that are expressed in the eye and brain. These genes include both well-studied and uncharacterized genes. For example, *miR181a* is expressed specifically in zebrafish RGCs later in development (Wienholds, 2005) and has a role in mammalian cellular differentiation (Judson et al., 2013), but its role is unknown within the context of eye development. Also near the top of the list (Table 4.1) is an equally intriguing but less mysterious gene, *nppa*. We found this gene interesting because it is a heart-specific gene whose expression is nearly undetectable in the RPCs (TPM=0.32) and in whole embryonic eyes, but it was identified

as a top candidate gene with ectopic expression in the retina when tet enzymatic activity is absent in *tet2*$^{-/-}$*;tet3*$^{-/-}$ mutant embryos as shown in our previous work (Chapter 3) (Seritrakul and Gross, 2017). This suggests that *nppa* is perhaps among the top epigenetically-regulated genes whose expression is the most sensitive to changes in 5mC/5hmC levels.

In the comparison between 22hpf and 27hpf RPCs, we found that 5mC/5hmC profiles are essentially identical at candidate retinal genes (*vsx2, atoh7, crx, ptf1a*). This is a surprising finding given that 22-27hpf is a critical time window leading up to the first wave of differentiation events when RPCs give rise to RGCs and other early-born retinal neurons (Almeida et al., 2014; Hu and Easter, 1999; Schmitt and Dowling, 1996), and one would expect to see evidence of epigenomic modification when RPCs are undergoing reprogramming. Three plausible and not mutually exclusive scenarios may explain these results.

First, based on a close gene-by-gene inspection of epigenomic profiles in human and mouse whole embryonic retinae from the Pecan database (Aldiri et al., 2017), there appear to be similar DNA methylation profiles (WGBS tracks) at these same retinal specification genes, and the profiles do not change over the course of early retinal neurogenesis. Additionally, by layering on the chromatin marks from the same group of samples, it becomes clear that the TSS of these genes also possess consistent chromatin states for a bivalent domain (H3K4me1-3, H3K27me3, H3K9Ac), suggesting that DNA methylation and hydroxymethylation act cooperatively with other epigenetic mechanisms, namely histone methylation and acetylation, to turn the genes on during neurogenesis. This is consistent with recent findings that 5mC/5hmC profiles benefit from being analyzed in conjunction with other epigenomic marks that influence the

chromatin state in order to make meaningful interpretation possible (Mo et al., 2016; Ueno et al., 2016; Wang et al., 2018).

Second, our candidate gene analysis centers around the profile of known candidate retinal developmental genes and expands to surrounding regions. However, these may or may not be the true regulatory regions driving the expression of these genes. It is known that enhancers can reside far away from the genes that they are driving, and their locations may not be immediately identifiable (Bulger and Groudine, 2011; Zentner and Scacheri, 2012). Although many putative enhancers have been identified across many mammalian cell types (Lizio et al., 2015), their genomic coordinates in zebrafish are less known. Therefore, the locations of differential methylation and hydroxymethylation important for driving the expression of retinal neurogenesis genes may be outside of the windows we examined, and to identify them, additional information from other chromatin marks is required.

Third, it is possible that the pools of vsx2+ RPCs we isolated from the retinae contain a heterogeneous mix of RPC subpopulations with various epigenomic profiles, rather than a homogeneous pool of identical RPCs, and the signals are averaged out from bulk BS/OXBS profiling. This is more likely to be true especially for the later RPCs isolated at 27hpf shortly before the onset of differentiation. Recent findings suggest that what we thought of as a uniform, seemingly indistinguishable population of progenitor cells may in fact be composed of distinct subpopulations, each one with a unique potential for subsequent differentiation (Cepko, 2014; Johnson et al., 2015), and these progenitors could undergo a series of transformations through "transitional progenitor" steps (Jin, 2016). In this scenario, one could imagine an immature retina composed of a progenitor pool for early-born RGCs and a larger, co-existing pool for later-born cell types (MG, BP, rods). At 27hpf, the RGC progenitors begin to modify their epigenomic

and transcriptional profiles to differentiate into mature RGCs, and at the same time the late-born progenitors undergo epigenomic changes in a different direction. Because there are only a few RGCs born at 28hpf, their unique epigenetic signal is drowned out by the undifferentiated RPCs and/or the late-born progenitor pool. To directly answer this, we need to employ a modified strategy where each RPC is profiled uniquely for the epigenome and transcriptome. Recent development in the field of genome editing and single-cell sequencing now make this a more feasible approach, as outlined in Future Directions (Chapter 5).

Although we present here the first successful combined BS/OXBS analysis in the developing retina, several technical improvements could be made. First, increasing sequencing depth would greatly improve the data quality, as sequencing depth directly influences the confidence of 5mC/5hmC methylation levels, both at the initial methylation calls and later during the maximum-likelihood 5hmC estimation. In the case of a limited budget, enhanced reduced-representation methods (eRRBS and RR-OXBS) could be utilized to allow higher coverage on a subset of CpG islands around MspI (or another restriction enzyme of choice), although this will reduce the coverage in other parts of the genome (Garrett-Bakelman et al., 2015). Second, an independent, direct 5hmC profiling technique could be employed in conjunction with the indirect OXBS-based profiling when maximum confidence in 5mC/5hmC levels is required. The raw data from all three libraries (BS, OXBS, TAB-seq) can be combined and processed, again using the compatible maximum-likelihood program such as MLML (Qu et al., 2013). Finally, to generate a biologically meaningful set of methylome and hydroxymethylome profiles, many more samples need to be purified and sequenced. The technical limitation here is the limited availability of transgenic lines that express fluorescent reporter in only one cell type (*vsx2* in RPCs, *atoh7* in RGCs, for example). Fortunately, efforts in

enhancer trapping have made available many more transgenic lines with unique expression in various cell types, several of which are within the retina (Kawakami et al., 2010). Additionally, Crispr-Cas9 editing strategy now allows insertion of a Gal4 sequence directly into a GFP or mCherry line, allowing conversion of existing lines to be utilized in cell sorting. These recent developments will allow us to generate epigenomic profiles of all differentiated retinal cell types in the near future.

Chapter 5: Summary and Future Directions

5.1 Summary of Work

DNA methylation (5mC) and hydroxymethylation (5hmC) have been subjects of investigation over the past decade in both developmental and disease contexts, due to their regulatory roles in modulating the expression of diverse sets of genes in conjunction with other chromatin modifications. However, little is known regarding the genomic distribution of these epigenetic marks and the roles they play during eye development. My dissertation work directly aimed at profiling the genome-wide distribution of 5mC/5hmC and investigating the expression and function of both families of DNA methylation machinery: *de novo* DNA methyltransferases (dnmt3 family, which add 5mC to unmethylated C) and cytosine dioxygenases (tet family, which convert 5mC to 5hmC).

In Chapter 2, I showed that all members of the dnmt3 family are expressed in distinct and overlapping domains of the developing eye. However, when these genes are mutated, either individually or in pairs, the resulting embryos still develop normally without any overt phenotypes, suggesting a high level of functional redundancy (Seritrakul and Gross, 2014). In Chapter 3, I generated mutants for *tet2* and *tet3*, which are both expressed during retinal neurogenesis, and identified cell type-specific neuronal defects in the $tet2^{-/-};tet3^{-/-}$ double mutant embryos (Seritrakul and Gross, 2017). I found that the mutant retinal cells are specified but most failed to terminally differentiate, due to altered expression of genes involved in extrinsic signaling pathways. In Chapter 4, I generated the first genome-wide nucleotide-resolution maps of 5mC/5hmC distribution of retinal progenitor cells during neurogenesis, which show expected methylation patterns at

candidate retinal genes and enabled discovery of several novel genes and pathways potentially involved in retinal neuron differentiation.

These results lay the foundation for future investigation into more mechanistic details of DNA methylation and hydroxymethylation during retinal neurogenesis, using recently developed genome engineering and single-cell epigenomic profiling, as described below:

5.2 FUTURE DIRECTIONS

5.2.1 Functional test of dnmt3 tissue-specific requirements

In zebrafish, six members of the dnmt3 family exists, all with predicted *de novo* methyltransferase domains (Goll and Halpern, 2011). Results from Chapter 3 suggest that several of these enzymes function redundantly (dnmt4/dnmt7 and dnmt6/dnmt8, at the very minimum), and this resulted in the lack of phenotype at both the whole embryo level and eye-specific level. While it is feasible to generate combinatorial mutant lines that lack all genes in the dnmt3 family, this process is time-consuming, laborious and may or may not yield interpretable phenotypic results. Nevertheless our lab has generated a compound mutant for *dnmt3,4,5,6,8*, which still showed normal development and no eye phenotype (Angileri et al., unpublished observations).

As an alternative, a cell-specific targeted inactivation of *de novo* methylation is a more viable option. Over the course of time between the generation of my double mutants and the conclusion of this Dissertation, a mutation (R882H) in Dnmt3A has been identified in human patients and functionally validated in a mouse model for acute myeloid leukemia (AML). This R882H substitution is within the catalytic domain and causes the mutant protein to act in a dominant-negative manner, inhibiting the formation

of the Dnmt3A and Dnmt3B tetramers required for methyltransferase activity, and thus inhibiting the endogenous wildtype proteins' function (Kim et al., 2013a; Russler-Germain et al., 2014). This mutation will be useful in enabling functional testing of dnmt3-family genes without having to generate a compound mutant. This can be achieved using three approaches:

First, for investigating global effects, one would simply overexpress the mutated dominant-negative Dnmt3a^{R882H} in zebrafish, either by mRNA injection or transgenically under control of a global or heat-inducible promoter (e.g. Hsp70l:gal4; 10xUAS: GFP-Dnmt3a-DN). Second, to achieve more precise control, several existing driver lines could be used to activate Gal4 in specific cell types of interest, such as *vsx2*:gal4 (for RPC), *atoh7*:gal4 and *isl2b*:gal4 (for early and late RGC, respectively), and *opn1sw1*:gal4 (for PR). Third, using information from the methylome map generated in Chapter 5, one could target the engineered mutant protein to genomic loci that normally gain 5mC (and presumably lose transcriptional activity) during differentiation and assess the resulting phenotype when *de novo* methylation is inhibited. This is now feasible using the Crispr-Cas SunTag-directed targeting strategy, which has already been validated to function with the wildtype Dnmt3A (Huang et al., 2017). Using sgRNAs targeted to 5mC-gaining sites, the SunTag on the deactivated Cas9 can 'tow' and concentrate Dnmt3 (mutant or wildtype) protein to virtually anywhere in the genome. The resulting cells can then be sorted and analyzed for methylation and gene expression changes.

5.2.2 Tissue-specific analysis of Tet function

Results from gene expression and pharmacological rescue experiments in *tet2*$^{-/-}$*;tet3*$^{-/-}$ embryos indicate that tet enzymes functions are mediated non-autonomously via

cell-extrinsic pathways during retinal neuron differentiation (Chapter 3). An obvious question remains to be answered: what cells or tissues are the source of these extrinsic molecules? While mosaic analysis using shield-stage transplantation offers one way to address this, it is a relatively invasive, laborious process and does not allow much control of the specific number and location of the resulting clones. To avoid this, or as an alternative approach, I outline below two options that are available and offer more spatio-temporal control.

First, to generate tet2/3 null clones and simultaneously lineage-trace them in the retina, one could utilize the 2C-Cas9 strategy (Donato et al., 2016) where *tet3* mutation is induced mosaically on a *tet2*$^{-/-}$ background (phenotypically wildtype) by injecting a construct containing UAS:Cas9-Cre and sgRNAs targeting *tet3* under control of retina-specific Gal4 driver (e.g. *vsx2*:gal4) (Fig 5.1). This will result in a variable number of *tet2*$^{-/-}$*;tet3*$^{-/-}$ mutant clones among phenotypically wildtype *tet2*$^{-/-}$*;tet3*$^{+/?}$ neighbors and allow analysis of retinal neuron type-specific differentiation using conventional immuno-histological staining. Additionally, labeled (i.e. double mutant) cells can then be FAC-sorted for transcriptome and methylome analysis and mutation validation by PCR, in comparison to the neighboring non-GFP, phenotypically wildtype cells.

Second, to pinpoint potential sources and timing of the extrinsic signals required for neuronal differentiation, Tet2 or Tet3 overexpression constructs utilized in Chapter 3 can be driven using a cell-specific UAS driver in the *tet2*$^{-/-}$*;tet3*$^{-/-}$ background. Several UAS lines exist that will allow expression of the construct in both early in development (*vsx2*:gal4 for RPCs, *tfec*:gal4 for RPE; *cryaa*:gal4 for lens) and later in development (*isl2b*:gal4 for RGCs, *gfap*:gal4 for MG, *opn1sw1*:gal4 for cones, for example) in a cell type-specific manner. Any domains where Tet2 or Tet3 wildtype expression 'rescues' retinal neuron differentiation in the *tet2*$^{-/-}$*;tet3*$^{-/-}$ background is a candidate source of the

extrinsic signals mediated by tet enzymes. Tet2 and Tet3 catalytically in active constructs can be used as a control, which would result in a lack of rescued phenotype, and the embryos would look identical to those with the *tet2*$^{-/-}$;*tet3*$^{-/-}$ background.

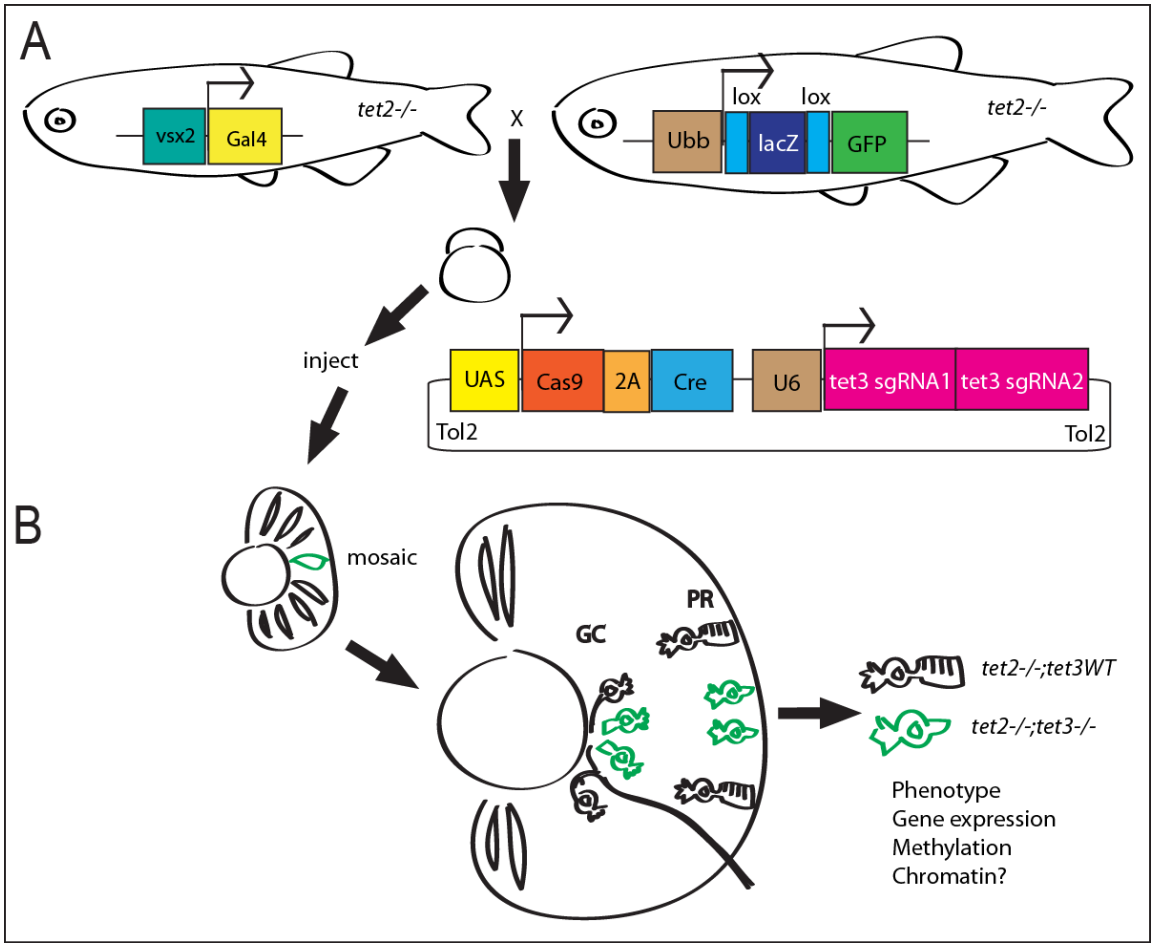

Figure 5.1: Schematic for clonal analysis of Tet loss-of-function experiments. See text for details. 2C-Cas9 construct adapted from (Donato et al., 2016).

5.2.3 Methylome and hydroxymethylome data mining

My analyses in Chapter 4 represent the tip of the iceberg in terms of the amount of information we can potentially learn from the BS and OXBS data generated, the additional data currently being generated (27hpf RPC RNAseq and replicates of BS/OXBS) and soon to be generated in the near future (BS/OXBS and RNAseq for RGCs and later born retinal neurons). As more bioinformatic tools and approaches become available, additional data mining is possible. For example, to streamline the analysis and reduce computational time, all of the methylation (5mC, 5hmC, C) levels were calculated from cytosines located within the CpG context, which is the most prominent form of methylation in vertebrate genomes (Smith and Meissner, 2013). However, DNA methylation is known to occur in a non-CpG context (CH = CA, CT, CC) at minimal levels in vertebrates, and at higher levels in plants (Zemach et al., 2010). A recent study reported that in postmitotic neurons, the combined CH level is typically less than 2%, and 5hmC occurs mostly in the CA context (5hmCA) near enhancer shores (Mellén et al., 2017). An advantage of the OXBS strategy we used is that methylation marks in all contexts are processed and sequenced at the same time. Therefore, additional methylation extraction rounds can be done on our existing OXBS reads, using Bismark or MethPipe, with modification to extract methylation score in a non-CG context, and subsequent downstream analysis can be performed as we did for CpG. This will yield the first non-CG methylome and hydroxymethylome maps in developing RPCs.

Additionally, regions of the vertebrate genome can be classified as partially methylated domains (PMDs), continuous regions where methylation levels are lower than genome average but not as low as those located within the hypomethylated regions (HMRs). These PMDs are often associated with gene repression and inactive chromatin marks, and genes located within PMDs show tissue-specific expression and are subjected

to the gain/loss of 5mC over the course of development (Gaidatzis et al., 2014; Schroeder et al., 2013). In practice, we can identify these PMDs in BS/OXBS data, compare methylation levels between 22hpf and 27hpf RPCs, and generate a list of differentially partially methylated domains (DPMDs). To identify biologically relevant genes within these domains, additional chromatin state data are likely required. This can be obtained by performing additional ChIP-seq for active/inactive histone mark and/or ATAC-seq. Alternatively, existing chromatin state data can be compared between mouse and human at an equivalent developmental time points (Aldiri et al., 2017). Although these data were generated from whole dissected retina, and not FAC-sorted RPCs, they should be comparable at an early developmental time because most of the retina is composed of RPCs.

5.2.4 Functional analysis of novel retina neurogenesis genes

The combined methylome and hydroxymethylome analysis in Chapter 4 identified a number of genes located near regions of differential methylation and hydroxymethylation (DMRs/DhMRs), some of which are known eye genes while several others are uncharacterized, potentially novel neurogenesis genes. The next step in validating these candidates is to perform a series of expression analyses by *in situ* hybridization at early (22hpf) and late (27hpf) time points. Utilizing the wildtype, tet double mutant line (*tet2*$^{-/-}$;*tet3*$^{-/-}$), and dnmt3 loss-of-function line (either a combinatorial knockout or the overexpressed dnmt3-DN as described above), one can begin to distinguish whether the transcription of these genes is indeed regulated by DNA methylation and/or hydroxymethylation.

According to the model presented in Chapter 3, one would expect to see cell type-specific expression patterns for 5mC-losing and/or 5hmC-gaining genes. For example, *slc17a6b* is not expressed in RPCs at 22hpf but strongly expressed in the inner retina, presumably RGCs, by the end of neurogenesis (Thisse and Thisse, 2004). In *tet2*$^{-/-}$*;tet3*$^{-/-}$, the expression would likely be absent because the gene stays silenced due to the lack of 5mC → 5hmC conversion (Fig 5.2A). Conversely, genes that gain 5mC likely turn from active in early RPCs to become repressed as RPCs differentiate into neurons, and eventually only maintain the expression in the CMZ where cells stay proliferative. In dnmt3 LOF embryos, this expression is likely expanded because 5mC could not be added *de novo* to silence the gene, and expression would be maintained in the specified and/or differentiated neurons (Fig 5.2B). This may also indicate that the cells, though specified, are not differentiated and remain ectopically proliferative, similar to the phenotype characterized in *tet2*$^{-/-}$*;tet3*$^{-/-}$. Lastly, for novel genes that show promising expression patterns, knockout lines can be easily and quickly generated using Crispr-cas9 to further functionally test their roles during retinal development *in vivo*.

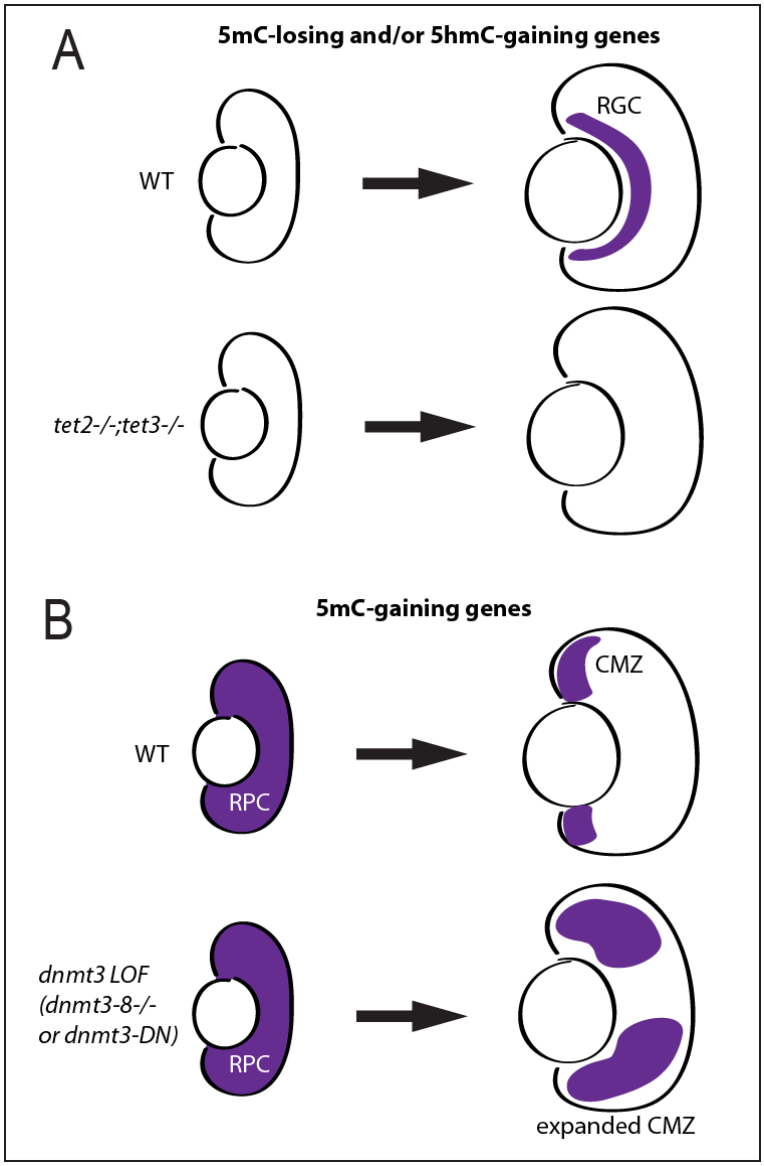

Figure 5.2: Expected expression patterns of candidate genes identified from OXBS analysis in embryonic retina. 5mC-losing and/or 5hmC-gaining genes likely become expressed in differentiated neurons, and this expression is abolished in *tet2*$^{-/-}$;*tet3*$^{-/-}$(A). 5mC-gaining genes likely lose expression upon differentiation, only staying on in the CMZ, and this domain is expanded in dnmt3 loss-of-function embryos.

5.2.5 Simultaneous lineage tracing and profiling of developing RPCs

The exact trajectory of retinal cell differentiation from RPCs to neurons and MG is still unclear. This is in contrast to the resolved cellular lineage in other systems such as hematopoiesis where multipotent hematopoietic stem cells give rise to various blood cell types in a clearly defined differentiation hierarchy (Rieger and Schroeder, 2012). Studies using long-term live imaging are beginning to document the stochastic behavior of differentiating RPCs both *in vivo* (He et al., 2012) and *in vitro* (Gomes et al., 2011), while functional experiments combined with mathematical modeling have elucidated the probabilistic and independent activity of core TFs in determining the fate of differentiating RPCs (Boije et al., 2015). These data support the stochastic model and imply that such a clearly defined lineage tree may not be necessary.

In contrast, experimental data show that RPCs undergo transitions from being multipotent to being restricted in differentiation potential as they turn on/off lineage-specific genes. For example, all early RPCs express *vsx2*, which acts to repress differentiation genes, and this expression decreases in most differentiating RPCs (Vitorino et al., 2009). During neurogenesis some RPCs turn on *atoh7* to become RGCs or PRs (Poggi et al., 2005), others turn on *vsx1* to become BPs or ACs (Jusuf et al., 2011), and those that retain *vsx2* expression are destined to become MGs and BPs (Vitorino et al., 2009). These data support the competency model and, at least to some degree, imply the existence of a lineage tree.

The seemingly indistinguishable RPCs are thought to be intrinsically unique (Cepko, 2014; Dyer and Cepko, 2001; Trimarchi et al., 2008) and there may exist "transitional" progenitors with dynamic epigenome and transcription profiles (Jin, 2016). In addition, single-cell RNAseq suggests that neural progenitors in the human cortex are indeed transcriptionally heterogeneous (Johnson et al., 2015). In Chapter 4, we observed

that the methylome and hydroxymethylome profiles are overall quite similar between 22hpf and 27hpf RPCs with only a few hundred DMRs and DhMRs across the genome, which may reflect, at least in part, the inherently heterogeneous nature of the RPC pools.

To resolve the retinal cell differentiation lineage and simultaneously profile the epigenome and transcriptome of individually developing RPCs and differentiated retinal cells, one could utilize the recently developed single-cell combined barcoding and epigenomic profiling approach. This can be done using the genome editing of synthetic target arrays for lineage tracing (GESTALT) (McKenna et al., 2016) by "barcoding" the early RPCs (22hpf) before differentiation and reconstructing the lineage relationship between differentiated cells at the end of neurogenesis (72hpf) based on the mutations accumulated within the barcodes. The cells carrying the barcodes can then be simultaneously profiled for DNA methylation and transcription using scNMT (Clark et al., 2018; Ren and Pott, 2017).

In practice, this would require generating a zebrafish line carrying a Cas9 construct under a heat shock promoter, a constitutive expression cassette with sgRNAs targeting the sequences on the barcodes, and a heatshock-inducible barcode array tagged with mCherry to allow for transcription (for RNAseq) and isolation of the labeled cells by FAC-sorting (Fig 5.3). The early RPCs will be heat shocked, and this triggers the Crispr-mediated mutation accumulation on the barcode over time. At the completion of neurogenesis, labeled cells will be isolated, and mRNA and genomic DNA extracted. The barcodes embedded within the mCherry mRNA will be sequenced along with the rest of the transcriptome, and a lineage tree will be constructed based on the "evolutionary" history of each cell. Meanwhile, the genomic DNA of each cell will be bisulfite treated and sequenced to obtain single-cell methylome profile.

It is important to note that this single-cell technology does not allow simultaneous profiling of 5hmC, as each strand of DNA can physically only be probed once for either 5mC or 5hmC, and not both. Additionally, although single-cell profiling is possible for 5hmC, it still relies on the enzymatic activity of BGT, and not an OXBS-based approach, thus will be limited in resolution (Mooijman et al., 2016). Regardless of the limitation, this combined approach (scGESTALT + scNMT) would allow tracing of the developmental trajectory of each retinal cell during neurogenesis and offer comprehensive methylome and transcriptome profile of those same cells. Additionally, depending on the topology of the tree, this approach may also lead to the discovery of new intermediate RPC types, and the transcriptome profiles may uncover novel genes and pathways previously undetectable by traditional bulk profiling methods.

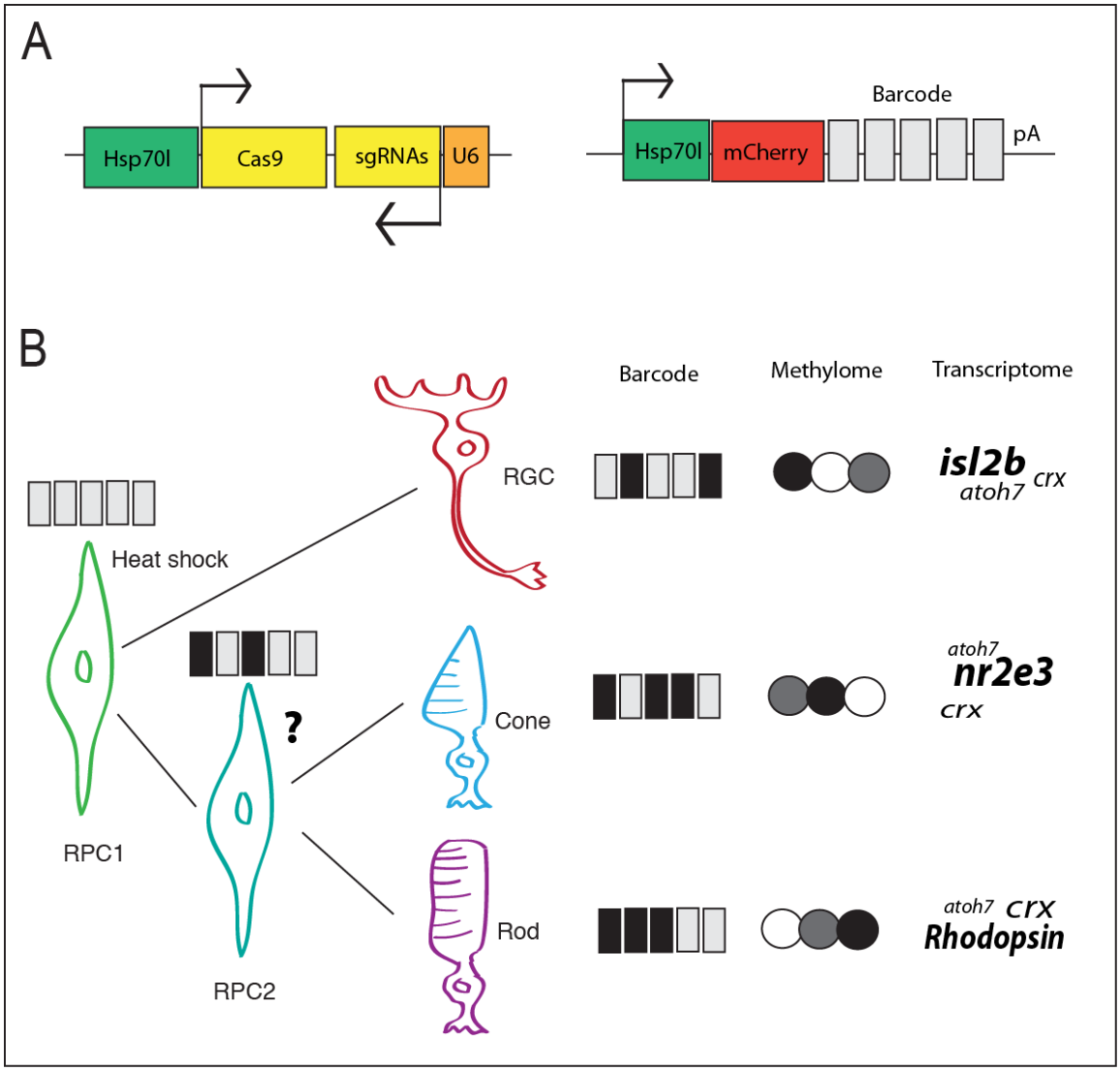

Figure 5.3: Schematic diagram of the combined scGESTALT + scNMT profiling approach in the developing retina. See text for details.

APPENDICES

Appendix A: Material and Methods

A1: Animal Lines and Husbandry

Zebrafish were maintained at 28.5°C on a 14/10 light/dark cycle and treated in accordance with the University of Texas at Austin and University of Pittsburgh IACUC regulations governing animal research. Euthanasia utilized tricaine, following procedures standard in the field and as approved by the IACUC. Embryos were incubated in the dark at 28.5°C and staged according to (Kimmel et al., 1995). Lines utilized in this study are: *tet2*^{au59}, *tet3*^{au60}, 10xUAS-nls-mCherry-t2a-myc-flag-TET2wt^{au50}, 10xUAS-nls-mCherry-t2a-myc-flag-TET2mut^{au51}, *Tg(isl2b:*GFP*)* (Pittman et al., 2008) and *Tg(vsx2*:GFP) (Vitorino et al., 2009).

A2: Phylogenetic analysis of proteins within dnmt3 and tet family

For dnmt3-family proteins, All amino acid sequences were obtained from NCBI with the following accession numbers: AAI63893.1 (zebrafish dnmt1); NP_001018153.1 (zebrafish dnmt2); NP_571461.1 (zebrafish dnmt3); AAI24099.1 (zebrafish dnmt4); NP_001018315.1 (zebrafish dnmt5); AAI62582.1 (zebrafish dnmt6); AAI63546.1 (zebrafish dnmt7); NP_001018144.1 (zebrafish dnmt8); AAH53047.1 (mouse Dnmt1); NP_001258682.1 (mouse Dnmt3A); NP_001258673.1 (mouse Dnmt3B); AAI26228.1 (human DNMT1); AAH 23612.1 (human DNMT3A); NP_ 008823.1 (human DNMT3B).

For tet-family proteins, Amino acid sequences were downloaded from NCBI, using the following accession numbers: NP_085128.2 (human TET1), NP_001240786.1 (mouse Tet1), AHE93329.1 (zebrafish tet1), NP_001120680.1 (human TET2), NP_001035490.2 (mouse Tet2), AHE93330.1 (zebrafish tet2), NP_001274420.1 (human TET3), NP_898961.2 (mouse Tet3), AHE93331.1 (zebrafish tet3). Alignments and phylogenetic trees were constructed using Geneious Tree Builder software with standard neighbor-joining method (Biomatters).

A3: IN SITU HYBRIDIZATION

Whole-mount in situ hybridization was performed essentially as described (Jowett and Lettice, 1994) with the following modification to allow probes to access the lens: (1) Full-length RNA probes were hydrolyzed with a mixture of 0.6 M sodium carbonate and 0.4 M sodium bicarbonate for 20 min. (2) Embryos older than 72 hpf were pretreated with 1 mg/ml Collagenase type IA before proteinase treatment step. DIG-labeled RNA probes for *notch1a, deltaA, ascl1a, vsx2, pax6a, neurod4* and *atoh7* were described previously (Uribe and Gross, 2010). Probes for *dnmt3-8*, *tet1*, *tet2, tet3, lef1, wnt1, wnt9B, opn1sw1, opn1mw1,* and *nppa* were cloned from zebrafish cDNA using primers listed in (Table A1) . PCR fragments were cloned into pGEM-T-Easy vector (Promega), sequence verified, linearized, and transcribed using SP6 and T7 polymerases with DIG RNA labeling mix (Roche). Synthesized RNA probes were purified using RNeasy kit (Qiagen), mixed 1:200 with hybridization buffer (50% formamide, 5xSSC, 0.1%tween, 5mg/ml yeast tRNA, 50µg/ml heparin), and heated to 68°C before use. Once the appropriate level of staining is achieved, embryos were re-fixed and embedded in tissue-freezing medium and cryosectioned at thickness of 12 mm.

primer name	locus	sequence
dnmt3 F	dnmt3	TGAATCAACACCATTTCCACGG
dnmt3 R	dnmt3	AAAGCCCAACAAGAGCACACTG
dnmt4 F	dnmt4	CGGCACCTTCAATCTCTGCTAT
dnmt4 R	dnmt4	GCCAGGGGGAAAACACAATACTC
dnmt5 F	dnmt5	TCCCACTGTTATGGAGAATGGTTG
dnmt5 R	dnmt5	TCCTGAAAATCACAAAGCGGC
dnmt6 F	dnmt6	TGGGGCAGGAAAAAAAAGTGAC
dnmt6 R	dnmt6	TCTCATAGGGCTCCGATTCAGG
dnmt7 F	dnmt7	AATCCTGATGACAAGTGCAGTCG
dnmt7 R	dnmt7	CAGCAAGTACCACATGACAGTTGG
dnmt8 F	dnmt8	ACACACACACAGCGACCTAAGAGG
dnmt8 R	dnmt8	CACAGTGAGGGAACGGTTTACCTG
opn1mw1_F1	opn1mw1	TGAACGGGACAGAAGGGAGC
opn1mw1_R1	opn1mw1	AAGTCTTGGAGAAGAAGGCTGGAA
opn1sw1_F1	opn1sw1	ACGTTGTGTGCGATGGAAGCG
opn1sw1_R1	opn1sw1	GACATAGGTGTACCAGTTCTGCACA
nppa_F1	nppa	CAGAGACACTCAGAGATGGCCG
nppa_R1	nppa	CCATTTAATGTTACTTACTGAAGGCAAAGGTC
tet1_F1	tet1	CTCAGGTATCAAAAACCTACGATGTGC
tet1_R1	tet1	TCCAGCCTCCTCTTGTGTGC
tet2_F1	tet2	CTTGAAACAAGGGCAGCCTTGT
tet2_R1	tet2	CTCATGAGTCTGTTGCTGAAGCAAG
tet3_F1	tet3	CAACCACCAACCTTCACCTT
tet3_R1	tet3	ATGCATGCAGTGATTCCTGA
wnt1 F	wnt1	ACTGTCGCATCAGAGATGCAGC
wnt1 R	wnt1	TATTTACCACCCCGGATTCAGAACA
wnt9B F	wnt9B	GTTTTAAAGAGACGGCCTTCCTGC
wnt9B R	wnt9B	CCCACTCATTTGCCTCTGTTTACTTC
lef1 F	lef1	ACACCACCCACAAGATGTCA
lef1 R	lef1	TGAACACACAGGTGCCAAAT

Table A1: List of primers used for cloning of in situ hybridization probes

A4: TALEN-MEDIATED GENOME EDITING

TALEN site were selected using a target selection tool, Mojo Hand, to identify suitable DNA binding sites (Neff et al., 2013). To facilitate identification and genotyping, TALENs were designed to cut at restriction endonuclease recognition sites. Plasmids encoding the repeat variable domain (RVDs), FokI nuclease domain, and other accessory sequences were assembled using the Golden Gate platform (Bedell et al., 2012b; Cermak et al., 2011; Dahlem et al., 2012). Completed constructs were *in vitro* transcribed and injected into 1-cell stage embryos. Resulting embryos were analyzed for mutations at the target site by restriction fragment length polymorphism (RFLP) and raised to adulthood. Potential founders were then analyzed for germline transmission of desirable loss-of-function mutations and outcrossed to wildtype fish in order to establish mutant lines. All TALEN-induced stable mutant lines currently maintained in the lab carry validated mutations that disrupt reading frames and are predicted to cause premature stop codons. Heterozygous adults of each line were identified by fin-biopsy and incrossed to generate homozygous mutant embryos. RFLP genotyping primers and restriction enzymes for dnmt-3 family genes are listed in Table A2.

To generate the *tet2*$^{-/-}$ and *tet3*$^{-/-}$ mutant lines, TALEN constructs were generated using Golden Gate assembly (Bedell et al., 2012b; Cermak et al., 2011) targeting the following sequences (spacer in bold and restriction endonuclease recognition sites underlined):

CATCCCAGATGGAAT**GGATAGA**<u>**TTTAAA**</u>**CTC**AACTTCTGCTTCAAC for *tet2*^{um59};

GCTCTGGGAGATAAACT<u>**GTAC**</u>**AGAGAAGTCAC**AGAAACCATCACCAAAT for *tet3*^{au60}.

Embryos at the 1-cell stage were injected with *in vitro* synthesized (Ambion) mRNA encoding the TALEN constructs (left and right arms) targeting *tet2* and *tet3*, separately, and raised to adulthood. At breeding age, potential founders were screened for germline transmission of mutations by sperm genomic PCR, followed by whole amplicon Sanger sequencing. Genotyping primers are listed in Table A2). Founders with deletions that resulted in frameshifts and premature stop codons were outcrossed to wildtype

females and embryos reared to adult. Potential heterozygotes were then screened for the desired mutation by restriction fragment length polymorphism (RFLP) using DraI for *tet2* and RsaI for *tet3*. RFLP fragments were resolved on a 1% agarose gel and mutant fragments were detected by the resistance to DraI and/or RsaI digestion (Fig. S2). To generate the double mutant line, heterozygotes carrying *tet2* or *tet3* mutations were crossed, and offspring were screened for the presence of both mutations by RFLP. *tet2*$^{+/-}$*;tet3*$^{+/-}$ fish were then incrossed to obtain *tet2*$^{-/-}$*;tet3*$^{-/-}$ embryos. Because homozygous mutation in either *tet2* or *tet3* alone does not affect viability, the remaining embryos survived to adult and a normal Mendelian distribution was obtained (Fig. S2).

Gene	restriction enzyme	primer	sequence
dnmt4	MboII	F	ACGTAATCTCGACCGCAAGGGA
		R	AAGGCCAAGCCATGTTGAAGC
dnmt6	AluI	F	GGGCCGAACTGTGTAATATTCTTAGGA
		R	TCAGGCTGGAACGAACCACTC
dnmt7	AluI	F	CTCTGGGTATTGCTGACATGCATTATTC
		R	ACTACGCATAGAACTTAGTGTGCAAGG
dnmt8	EcoRV	F	GAAAATGGATCTAATGGCGTGTGTGATAT
		R	AGGAGGTCATCAGTTTTCAGGATTAAAAG
tet2	DraI	F	CACAAACCTCTCAGACAGGTCAGT
		R	TCTCTGTTGACTTTCAGGGGCAG
tet3	RsaI	F	CAATGCCTAGATCAACCACTTAGTGTC
		R	GTATCAGGAATGTGCAAACATCTCATTTG

Table A2: List of primers and restriction enzymes used for genotyping of the mutant lines.

A5: CRISPR-MEDIATED GENOME EDITING

CRISPR line for dnmt4 deletion was generated following the protocol as described (Auer et al., 2014). Oligo sequences for the construction of sgRNAs were GCAGCTAATACGACTCACTATA**GGTGAGTGTACCTGCGGGTG**GTTTTAGAG CTAGAAATAGCAAG (for targeting exon3) and GCAGCTAATACGACTCACTATA**GGCGCCAGACAGAAGCTGCT**GTTTTAGAG CTAGAAATAGCAAG (for targeting exon 22). Targeting sequences are in bold. Deletions were detected by PCR using primers E3_GTF1: CTGTCACACATCTTTCCCAGATGAC, E3_GTR1: CCATTGACTGAAAGCTTCTTCAGGGC for exon3 and E22_GTR2: TTGAGTTTGGCCTGTTGCTCAA. If the deletion is present, the genomic DNA is amplified with E3_GTF1 and E22_GTR2. However, if the deletion is absent, the PCR for E3-E22 fails, and genomic DNA is instead amplified with E3_GTF1 and E3_GTR1.

A6: GENERATION OF TET OVEREXPRESSION TRANSGENICS

10xUAS:mCherry-Tet2WT^{au50} and *10xUAS*:mCherry-Tet2mut^{au51} were generated using mouse Tet2 wild-type and catalytically inactive constructs (Ito et al., 2010). These domains were reconfigured into zebrafish inducible overexpression constructs using the Tol2 Kit (Kwan et al., 2007). Briefly, the entire coding sequences were amplified and ligated, along with a mCherry-T2A-myc coding fragment, into Gateway-compatible pME-MCS using Gibson assembly system (NEB). These intermediate fragments were then assembled into expression constructs consisting of p5E-10XUAS, pME-mCherry-T2A-Myc-Flag-Tet, p3E-poly-A-tail, and the pDestCG2 backbone carrying *cmlc2*:GFP. The final DNA constructs were injected into 1-cell stage zebrafish embryos along with *Tol2* mRNA for stable genomic integration (Kwan et al., 2007). Embryos were screened for the heart GFP expression at 3-4dpf and raised to adult. Multiple founders were screened for germline transmission and the strongest expressing lines were kept for experiments.

A7: RT-PCR AND WESTERN BLOT ANALYSIS

RT-PCR for *tet2* and *tet3* was performed using exon-spanning primers listed in Supplementary Table 3. Embryos at 2dpf and 5dpf (n=20 per genotype per condition) were euthanized and RNA extracted using RNeasy kit (Qiagen). cDNA libraries were generated using iScript cDNA synthesis kit (BioRad).

Western blot analyses was performed essentially as described (Tittle et al., 2011) with slight modifications. At 3dpf, 40 embryos per condition were euthanized, de-yolked, and protein extracted. Samples were separated by electrophoresis on 4-12% bis-tris gel with NuPage MOPS SDS running buffer (Invitrogen) and transferred to PVDF membrane at 30V for 2 hours, then at 12V overnight at 4°C. Membranes were blocked in 1%BSA, 5% non-fat milk in TBST for 2 hours at RT and incubated in anti-TET3 rabbit polyclonal antibody (ab139311, Abcam) overnight at 4°C, then washed, incubated with HRP-conjugated donkey-anti-rabbit antibody (711-035-152, Jackson Immuno Research), rinsed, and incubated with substrate solution (Super Signal West Femto, Thermo Fisher). Images were acquired and band intensity quantified using ChemiDoc XRS+ system (BioRad). For normalization, membranes were stripped for 12 minutes in Restore Western Blot stripping buffer (Thermo Fisher), rinsed, re-blocked, probed with anti-actin mouse monoclonal antibody (CP01, Millipore) followed by HRP horse-anti-mouse secondary (7076, Cell Signaling) and imaged as above.

A8: BRDU INCORPORATION, PLM ASSAYS, AND TUNEL ASSAY

5-bromo-2-deoxyuridine (BrdU) incorporation was performed using a 15-minute pulse for PLM assays, and a 2-hour time window for 48hpf-5dpf assays. Embryos were treated in 0.3% BrdU, fixed in 4% PFA in PBS, embedded, and cryosectioned at 12μm. Sections were treated with 4M HCl at 37°C for 10min, blocked in block solution (5% normal goat serum, 0.1% tween, 1% DMSO, in PBS), incubated with anti-BrdU (1:250; Abcam) in block overnight at 4°C, stained with anti-rat Cy3 secondary (1:250) and counterstained with Sytox green at 1:10,000 (Molecular Probes). Cells undergoing mitosis were detected using anti-phospho histone H3 (ser10) (1:250) (Millipore 06-570),

stained with anti-rabbit Cy2 secondary (1:250), and counterstained with DAPI (1:500). For TUNEL analysis, embryos at 36hpf, 3dpf, 4dpf, and 5dpf were fixed in 4% PFA 1xPBS at 4°C overnight, cryosectioned at 12μm, and processed for TUNEL using TMR Red in situ cell death detection kit (Roche) per manufacturer's protocol.

A9: IMMUNOHISTOLOGICAL ANALYSIS

Immunohistochemistry was perform as described (Uribe and Gross, 2007), with the following antibodies: zpr-1 (cones; ZIRC), zpr-3 (rods; ZIRC), zrf-1/gfap (Muller glia cells; ZIRC), Zn8 (ganglion cells; ZIRC), and HuC/D (ganglion and amacrine cells; Molecular Probes). Embryos were cryosectioned at 12μm and incubated with primary antibodies diluted at 1:200 in block overnight at 4°C, then incubated with secondary antibody (anti-mouse Cy3) for 2hrs. Sections were counterstained with Alexa Fluor-633 Phalloidin at 1:100 and Sytox green at 1:10,000 (Molecular Probes) or mounted using Vectashield with DAPI (Vector Labs).

A10: WHOLE-MOUNT CHROMOGENIC IMMUNOSTAINING

Embryos were fixed in 4% PFA 1x PBS at 4°C overnight, rinsed once in PBST (0.1% tween-20, 1xPBS), once in water, and treated with 100% acetone for 7min at -20 °C to permeabilize the tissue, then rinsed one time each in water, PBST and PBDTX (1%BSA, 1%DMSO, 0.1% TritonX, 1xPBS, pH=7.3). Embryos were blocked for 1hr (2%NGS in PBDTX), incubated in Zn8 primary antibody (ZIRC) at 1:200 dilution 4°C overnight, washed 4 x 20min in PBDTX, and incubated in secondary horse anti-mouse HRP-tagged secondary (Cell Signaling) at 1:1,000 dilution for 2hrs. Embryos were then washed in PBSTX (0.5% Triton-X, 1xPBS) 4 x 20min, incubated in DAB working solution (Vector Labs) for 2-10min until staining was visible, rinsed in water, and stored in PBS before imaging.

A11: RNA EXTRACTION AND TRANSCRIPTOME ANALYSIS

One hundred zebrafish eyes were dissected at either 36hpf or 72hpf using a flame-sharped tungsten wire, and RNA extracted using Qiagen RNeasy kit as described (Uribe et al., 2012). For 36hpf, library preparation with polyA mRNA capture and sequencing was performed using Illumina NextSeq 500 paired-end 75bp reads. 450 million reads were generated. Raw FASTQ sequences were quality checked, trimmed, and mapped using CLC Genomic Workbench 9.0.1 to zebrafish reference genome GRCz10 at 85% mapping efficiency. Transcript abundances were calculated and differentially expressed genes (DEGs) were identified using CLC Genomic Workbench 9.0.1. For 72hpf, library was prepared as above and sequenced on an Illumina HiSeq 2500 PE2x125. 66 million reads were generated. FASTQ sequences were quality checked using FastQC (Babraham Bioinformatics), mapped to GRCz10 using TopHat, and DEGs were identified using Cufflinks package from Tuxedo suite (Trapnell et al., 2012). Genes with expression values above log2 fold-change of 2 are considered differentially expressed. All computational analyses utilized the Texas Advanced Computing Center and University of Pittsburgh Center for Simulation and Modeling. Raw and processed data are publicly available through NCBI Gene Expression Omnibus (accession number GSE80134).

Functional annotation was done using DAVID Bioinformatics 6.8 (https://david.ncifcrf.gov). Differentially expressed gene lists from RNAseq were filtered for log2 fold-change of 2 or higher and submitted to DAVID Gene Ontology for biological pathways (GOTERM_BP_DIRECT).

A12: MOSAIC RETINAL ANALYSIS

Shield-stage transplantation experiments were performed essentially as described (Carmany-Rampey and Moens, 2006). Embryos were injected with Alexa Fluor 488 dextran (10,000 MW, anionic, fixable) diluted at 1% in 0.2M KCl. Cells were transplanted from labeled donor embryos into unlabeled host embryos at the shield stage, targeting the presumptive retinal field (Woo and Fraser, 1995). Approximately 10 cells were transplanted per host embryo to minimize the 'community' effect resulting from

clones that are too large. Embryos were sorted at 24hpf for donor clone contribution and fixed at 72hpf for sectioning and immunohistochemistry.

A13: IMAGING

Confocal imaging of cryosectioned and immuostained embryos was done using Zeiss LSM5 and Olympus FV1200 confocal microscopes. Whole-mount images were taken using Leica MZ16F stereomicroscope. Sectioned *in situ* hybridization images were captured using Leica DM2500. All images were analyzed using ImageJ (imagej.nih.gov). Cell counting was done using ImageJ Cell Counter plug-in.

A14: SITE-SPECIFIC 5MC AND 5HMC QUANTIFICATION

Bisulfite sequencing was performed using EZ DNA Methylation-Direct kit (Zymo Research) using primers listed in Table A3. Eye tissues at 72hpf were dissected (n=7 per condition) and immediately processed through proteinase K digestion and bisulfite conversion. Converted DNA was purified and amplified using hot-start ZymoTaq and bisulfite-specific primer pairs (Supplementary Table 3). PCR amplicons (~300bp) were either directly sequenced or sub-cloned for sequencing. Sequencing traces were analyzed using QUMA (RIKEN, Japan). Our bisulfite treatment procedure generally yielded ~98% conversion. Clones that contained low quality sequences were manually excluded from the analysis.

Locus-specific 5hmC quantification was performed using the Quest 5hmC Detection kit (Zymo Research) using primers listed in Table A3. Briefly, genomic DNA was extracted at 72hpf using Purelink Genomic DNA purification kit (Invitrogen). Genomic DNA was divided into three groups: 1) Glycosylated and digested with a glucosyl-5hmC sensitive endonuclease, MspI [+GT]; 2) Unglucosylated and digested with MspI [-GT] (negative control); 3) unprocessed genomic DNA [untreated] (positive control). All DNA samples were purified, and equal amounts used as templates for quantitative real-time PCR. Quantitative real-time PCR was done using SYBR green master mix in 10ul volume, and reactions run in a CFX384 detection system (BioRad). Cq values were first evaluated by comparing the difference between the +GT and -GT,

and between +GT and untreated. If Cq_{+GT} is close to (<1 Cq difference) $Cq_{untreated}$, the locus is considered fully hydroxymethylated. Conversely, if Cq_{+GT} is close to Cq_{-GT}, the locus is considered non-hydroxymethylated. For each locus with Cq values that pass these criteria, they are considered partially hydroxymethylated, and percent 5hmC was calculated as follows: $\%5hmC = \{[Cq_{-GT} - Cq_{+GT}] / [Cq_{-GT} - Cq_{untreated}]\} *100$.

Name	Purpose	Region	Sequence
sw_TSS_bsF1	bisulfite sequencing	opn1sw1 TSS	TAGGGAGATGTTTATTAAGATGTAGTTGAGAGG
sw_TSS_bsR1	bisulfite sequencing	opn1sw1 TSS	ATCCCCRCCTTTTCACACCAAACAAAAC
sw_TSS_bsF2	bisulfite sequencing	opn1sw1 TSS	GGTGGTTTTTTATAAAATTTTTAAAGTTTTGTTTGGTG
sw_TSS_bsR2	bisulfite sequencing	opn1sw1 TSS	CTTTAAAAAATCTTCTCCTATATTCAACAAAAACTC
sw_TSS_quF1	Quest-5hmC detection	opn1sw1 TSS	TGAAAAGGCGGGGATGAAAAGG
sw_TSS_quR1	Quest-5hmC detection	opn1sw1 TSS	CGACACGTGTGGCCTAATCAA
sw_TSS_quF2	Quest-5hmC detection	opn1sw1 TSS	AGTCAACAAGGAGCCTGACTGAA
sw_TSS_quR2	Quest-5hmC detection	opn1sw1 TSS	TCACATGGACAGGGCGGATTAT
mw_TSS_bsF1	bisulfite sequencing	opn1mw1 TSS	ATATAGGYGGTTTTAATTTAGGTTTTAGTAAAAATG
mw_TSS_bsR1	bisulfite sequencing	opn1mw1 TSS	CCAAAAACCATATAACCCACAAATAAAC
mw_TSS_bsF2	bisulfite sequencing	opn1mw1 TSS	ATTAATTYGTGTTGGAAGAAATGGATATTATGTTG
mw_TSS_bsR2	bisulfite sequencing	opn1mw1 TSS	AAACTTCRCACAAACCCAATCCTATTTAAC
mw_CCGG_bsF1	bisulfite sequencing	opn1mw1 MspI site	TAAAGTTGATATGTTAGTTTTTGTTATGTGAGTTGTAG
mw_CCGG_bsR1	bisulfite sequencing	opn1mw1 MspI site	TAAAAAATTACAATAAAAAAAAAACCCAATAAAAAACC
mw_CCGG_quF1	Quest-5hmC detection	opn1mw1 MspI site	CATGTGAGTTGCAGCGTACTTTCAG
mw_CCGG_quR1	Quest-5hmC detection	opn1mw1 MspI site	CGCAAATGACAGCAGCCTTTAGG
OpCGI_bsF1	bisulfite sequencing	opsin CpG island	TTTTTTTGGAYGTTTGAAGAGATTGTAGTGTTTG
OpCGI_bsR1	bisulfite sequencing	opsin CpG island	AAACCTCRACTAAAAAACTAAAACCTATC
OpCGI_bsF2	bisulfite sequencing	opsin CpG island	GTAAAGYGAGGGGGTATTAGTTGAAAAG
OpCGI_bsR2	bisulfite sequencing	opsin CpG island	ATAACRCAAACTAAACATCCTATAACC
OpCGI_quF1	Quest-5hmC detection	opsin CpG island	ATAGTGGAGGAAGAAAGCCGAGAG
OpCGI_quR1	Quest-5hmC detection	opsin CpG island	TCTCCACGTCCTCCTTCTCCT
OpCGI_quF2	Quest-5hmC detection	opsin CpG island	CGTAAAGCGAGGGGGTATTAGTTGA
OpCGI_quR2	Quest-5hmC detection	opsin CpG island	CGTGCTGCCGATTTTAACCGAA
nppa_TSS_quF1	Quest-5hmC detection	nppa TSS	CCTTCAGTCTGCATCATTGGCC
nppa_TSS_quR1	Quest-5hmC detection	nppa TSS	GAGAAGTCCTGGTAACTCGGACAC
nppa_TSS_quF2	Quest-5hmC detection	nppa TSS	ATGAGCCACTGACAGCCAACT
nppa_TSS_quR2	Quest-5hmC detection	nppa TSS	TGTCAGAGCTGTGTCTGCTCAA

(Continued on the next page)

nppa_TSS_bsF1	bisulfite sequencing	nppa TSS	TTATATAGTGTYGTAATAGTGTAGTAAAATATTATGTG
nppa_TSS_bsR1	bisulfite sequencing	nppa TSS	TAACTATTTTTCRCAAACCTAAAAAAATCCTAATAAC
nppa_TSS_bsF2	bisulfite sequencing	nppa TSS	TTGTTTGTGTTYGAGTTATTAGGATTTTTTTAGGTTTG
nppa_TSS_bsR2	bisulfite sequencing	nppa TSS	TTAATTAAACATAAACCCAAATACCTTAAACTTAACC
nppa_body_quF1	Quest-5hmC detection	nppa gene body	GGCATCAGAGAGAGCCGTAGATTATG
nppa_body_quR1	Quest-5hmC detection	nppa gene body	AGCCAGACAAGCTTTTGCTTCG
nppa_body_quF2	Quest-5hmC detection	nppa gene body	GCGCACCACTTTTGTCCAAAGT
nppa_body_quR2	Quest-5hmC detection	nppa gene body	GCTCTCAAGACTTACCTGATCTCGG
nppa_body_bsF1	bisulfite sequencing	nppa gene body	TTTATAAYGTTATGTTTTTTTTTAGAGTTTGTTGTAG
nppa_body_bsR1	bisulfite sequencing	nppa gene body	ATTACAACCRAAAATACTAAAAAACCCTATAC
nppa_body_bsF2	bisulfite sequencing	nppa gene body	YGGTAAATTTTTAAGTATGTAAGGGGG
nppa_body_bsR2	bisulfite sequencing	nppa gene body	AAACRAAAATATTACCTTATCTCAAACACAATAC

Table A3: List of primers used for bisulfite sequencing and Quest analysis

A15: ENZYME-LINKED IMMUNOSORBENT ASSAY (ELISA)

Sandwich-based 5hmC ELISA was performed according to the manufacturer's protocol (Zymo Research). Genomic DNA samples were extracted at 5dpf using Purelink Genomic DNA extraction kit (Invitrogen), and diluted to 1ng/µl in water, denatured by heating at 98°C and cooling on ice, and 5hmC DNA was bound to the ELISA plate coated with anti-5hmC polyclonal antibody (1:1,000). Bound DNA was detected with anti-DNA HRP antibody (1:100), and was allowed to develop for 20min. 410nm absorbance was measured by a plate reader (BioTek), and a standard curve generated using linear regression from five DNA samples with known concentrations of 5hmC. Percent 5hmC was calculated as follows: %5hmC = (absorbance - y-intercept)/slope. Note that percent 5hmC in ELISA is based on the total number of hydroxymethylated cytosines, calibrated to standards (set of DNA with known 5hmC%). For example, 0.1% 5hmC means 1 of every 1,000 cytosines is 5hmC. Percent 5hmC in the site-specific glucosylation/digestion (Quest) assay represents the relative amount of 'protected' 5hmC

at each MspI (CCGG) site analyzed, compared to the two internal controls for each locus: fully digested DNA (representing 0% 5hmC) and undigested DNA (representing 100% 5hmC). Thus, these two numbers are not directly comparable, but should be in agreement with each other.

A16: PHARMACOLOGICAL TREATMENT

tet2$^{-/-}$*;tet3*$^{-/-}$ mutants and sibling embryos carrying *isl2b*:GFP transgene were dechorionated and incubated from 24hpf to 72hpf in embryo medium with 50μM DAPT (N-[N-(3,5-difluorophenacetyl)-L-alanyl]- S-phenylglycine-t-butyl ester; InSolution γ-secretase inhibitor IX, 565784, Calbiochem), 5μM IWR-1-endo (5.04462.0001, Calbiochem), or 1% DMSO as vehicle control. For BIO treatment, wildtype embryos carrying *isl2b*:GFP transgene were incubated in 2μM BIO (2'Z,3'E-6-Bromoindirubin-3'-oxime, B1686-5MG, Sigma-Aldrich) from 24-72hpf. All embryos were fixed at 72hpf in 4%PFA 1xPBS, sectioned, and processed for immunostaining. Optic nerve diameter measurements were done in 5-7 embryos per condition at optic nerve head, using FluoView software (Olympus). P-values were calculated using two-way ANOVA with multiple comparison (for DAPT and IWR) and two-tailed unpaired t-test (for BIO) using Prism GraphPad.

A17: RPC ISOLATION BY FACS AND GENOMIC DNA EXTRACTION

Zebrafish embryos at 22hpf or 27hpf were euthanized in Tricaine solution on ice for 5 minutes. Whole eyes were dissected using a flame-polished tungsten needles, rinsed in cold PBS. Eye tissues were dissociated using 0.25% trypsin/EDTA for 5 minutes at room temperature, passed through a 70um cell strainer, and the reaction was stopped using 2mM $CaCl_2$ in 5%FBS in PBS. Cell suspension was then pelleted down using a centrifuge at 450g for 5 minutes. Cells were then re-suspended in PBS and stained using Live/Dead dye (1ul/mL) according to manufacturer's instruction (Thermo Fisher), washed once with PBS, and resuspended in final solution of 5%FBS/PBS. Cells were kept on ice until sorting, typically within two hours after dissection. Cell sorting was

done using FACSAria II machine with FACSDiva software. Gates were set to only allow live cells (negative for Live/Dead signal) and bright vsx2:GFP signal. Cells were then collected into tubes containing 5%FBS in PBS for processing. Genomic DNA was extracted using Quick-DNA plus kit (Zymo), and total RNA was extracted using RNeasy mini kit (Qiagen) per manufacturer's protocols.

A18: OXBS Library Preparation and Sequencing

After ensuring sufficient starting genomic DNA quantity of 200ng per reaction, samples were processed for OXBS library construction using TrueMethyl Whole Genome kit (Cambridge Epigenetix). Briefly, starting genomic DNA samples were spiked with sequencing and digestion control fragments containing 5mC and 5hmC at known sites, then denatured and purified using magnetic beads. For BS/OXBS treatment, the samples were divided equally into two halves: OXBS samples were treated using oxidizing reagent, while BS samples only undergo mock oxidation. Both samples were then processed in parallel through conventional bisulfite conversion reaction, desulfonation, and magnetic bead purification. For library construction, both BS and OXBS DNA were end activated, complementary strand synthesized, indexed and adapter ligated. The finished libraries were then processed for final quantification and normalization before loading into Illumina NextSeq 500 (High Output, 150 cycle) for deep sequencing.

A19: OXBS Data Processing and Alignment

Raw fastq files from the sequencing runs were trimmed and quality assessed using TrimGalore and FastQC (Babraham Bioinformatics). Trimmed sequences were aligned using BWA-meth (Pedersen et al., 2014) with default parameters to the zebrafish bisulfite genome (GRCz10), which was prepared to include both unconverted and fully converted bases. Mapped reads were then converted from SAM to BAM (Li et al., 2009), imported into .mr format for MethPipe (Song et al., 2014), sorted, and duplicated removed. Calculation of methylation level and extraction of methylation score was done using

bsrate (`bsrate -c`) and methcount (`methcounts -n -c`) respectively (Song et al., 2014). Note that during methcount step, -n and -c flags were specified to generate counts for only symmetric CpG and exclude methylation in non-CpG context to ensure data compatibility with subsequent analysis steps. Matched BS and OXBS methylation count files from each experiment were then processed using MLML (Qu et al., 2013) to estimate the 5mC and 5hmC levels at each nucleotide using a maximum-likelihood estimation method (`mlml -u BS.meth -m OX.meth -o 22_walt_cpg.mlml`). 5mC and 5hmC levels for all CpG across the genome were filtered to include only those with at least 1 read coverage for both BS and OXBS datasets (i.e. exclude any sites without supporting read evidence from both datasets), and this was output into pseudo methcount files (.meth) which function similarly to meth count files from a traditional bisulfite experiment.

A20: METHYLOME ANALYSIS AND ANNOTATION

Identification of regions with hypo-methylated 5mC or hyper-methylated 5hmC were done using hmr program within the MethPipe suite (Song et al., 2014). For hyper-5hmC, meth count files were first inverted and ran through hmr similarly to hypo-5mC files. Differential methylation scoring, identification of differentially methylated and hydroxymethylated regions (DMRs, DhMRs) were done using methdiff and dmr programs within MethPipe, respectively. DMRs and DhMRs were filtered using awk to include only site with a minimum of 5 total CpGs and 3 significant CpGs (significant cutoff value at default of p=0.05). Regions with intersecting DMRs and DhMRs were identified using bedtools intersect (Quinlan and Hall, 2010). For annotation of DMRs and DhMRs, region files in bed format were annotated using annotatePeak.pl in Homer software suite (Heinz et al., 2010) using default parameters against the zebrafish genome GRCz10 database.

Appendix B: *au8* and *occ* mutant mapping

au8 is a mutant presented with cataract lens phenotype identified from an ENU mutagenesis screen within the Gross lab (Lee et al., 2012). Similarly, *occ* is a mutant independently isolated by another forward genetic screen at University of Washington. Both mutants display strikingly similar lens phenotype where the lens capsule rupture and part of the lens cortex spills out of the capsule.

To identify causative mutations, *au8* and *occ* mutants were mapped using the NGS-based 'cloning by sequencing' approach. *au8* heterozygous carriers (AB background), identified by pair-wise incrossing, were outcrossed to wildtype TU fish (ZIRC) to generate AB/TU F1 hybrid mapping line. Heterozygous F1 fish were identified and incrossed to produce F2 homozygous mutant embryos. Genomic DNA was extracted from 48 F2 homozygous mutant embryos at 6dpf using DNeasy Blood and Tissue kit (Qiagen) and sent to next-generation sequencing on Illumina HiSeq 2000 (2x100 paired-end) at the University of Texas at Austin Genomic Sequencing and Analysis Facility.

For *occ* mutant, mapping line was generated as above, with the exception that the outcross was done by crossing *occ* carriers (AB background) to a TU/WIK mixed wildtype background. Mutants were identified, and genomic DNA extracted as above. DNA was sent to Vanderbuilt University Medical Center for sequencing on Illumina HiSeq 2500 (2x100 paired-end). Both lines were sequenced to an average genome depth coverage of 15X.

To identify putative mutation site, raw sequencing reads were analyzed at Texas Advanced Computing Center using the BSFseq (bulk segregant linkage) mapping

function of MegaMapper pipeline (Megason Lab, Harvard Medical School) (Obholzer et al., 2012). Both mutations were predicted to be in *col4a5* on chromosome 7.

au8 mutation was mapped to a C to T conversion (C52621050T) on exon 26, which generated a premature stop codon, truncating the protein at 665 amino acids (aa) out of 1,659 aa full-length protein [p.Gln665Stop]. Mutation was verified by genomic DNA PCR and Sanger sequencing. For complementation test, *au8* heterozygotes were crossed to *col4a5*^{sa1609}, a known *col4a5* mutation carrier from the Zebrafish Mutation Project (Kettleborough et al., 2013). Resulting mutant embryos (26.9% of total embryos) from *au8* x *col4a5*^{sa1609} displayed lens phenotype similar to *au8*, thus confirming that both mutations are in the same gene.

occ mutation was mapped to a 5' splice junction of intron 41-42 where a T to G conversion (T50904935G) is predicted to cause intron inclusion and a premature stop codon located within the intron. This resulted in a truncation of the protein at 1,315 aa out of 1,659 aa full-length [p.Ser1308GlnfsStop6]. Causative mutation was verified by Sanger sequencing of *occ* mutant genomic DNA, and the presence of intron 41-42 in spliced mRNA was verified by RT-PCR and Sanger sequencing of cDNA.

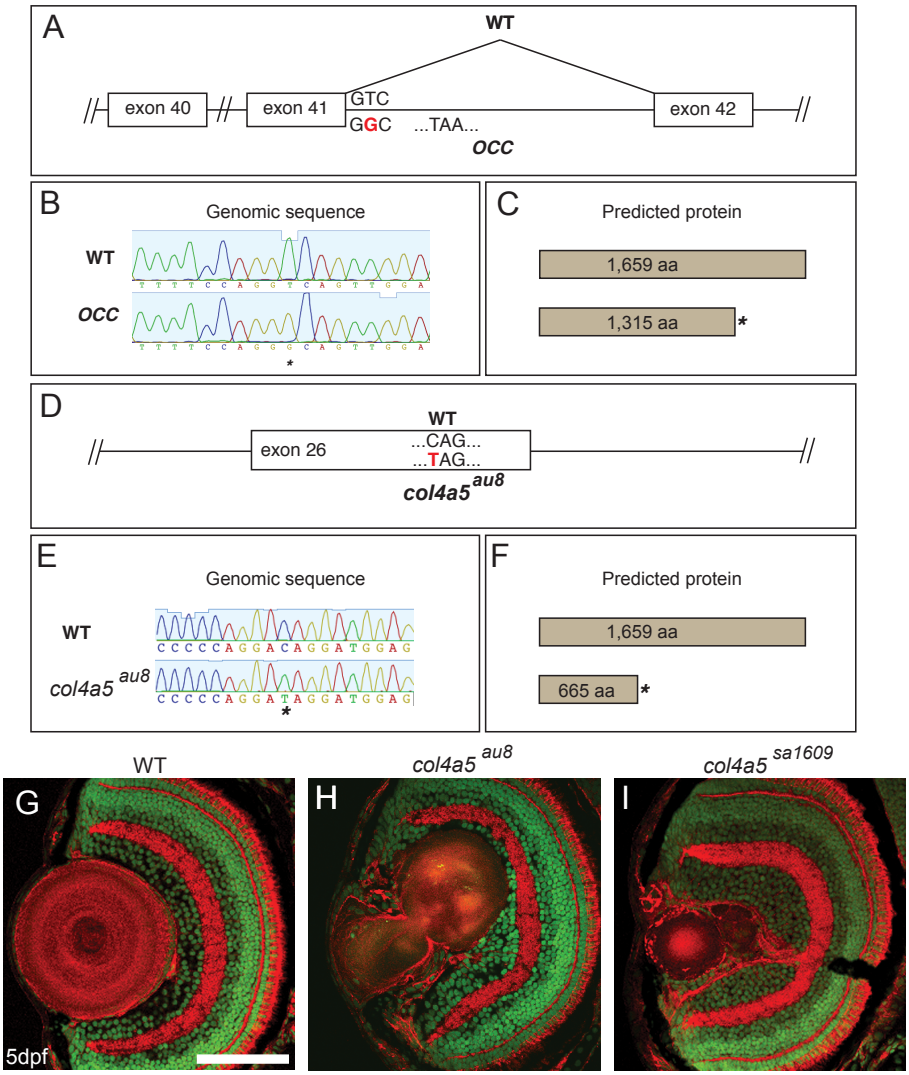

Figure B.1: Causative mutation site for occ mutant is located at the downstream splice junction of *col4a5* exon 41, which resulted in an intron inclusion and a premature stop codon within the retained intron (A-C), truncating the protein at 1,315aa. Causative mutation of *au8* is located within *col4a5* exon 26. The single nucleotide C→T substitution introduces a premature stop codon, truncating the protein at 665aa (D-F). Both mutant embryos display ruptured lens capsule and disrupted structural integrity of the lens cortex, compared to sibling, at 5dpf (G-I).

References

Agathocleous, M. and Harris, W. a (2009). From progenitors to differentiated cells in the vertebrate retina. *Annu. Rev. Cell Dev. Biol.* **25**, 45–69.

Albadri, S., Del Bene, F. and Revenu, C. (2017). Genome editing using CRISPR/Cas9-based knock-in approaches in zebrafish. *Methods*.

Aldiri, I., Moore, K. B., Hutcheson, D. a, Zhang, J. and Vetter, M. L. (2013). Polycomb repressive complex PRC2 regulates Xenopus retina development downstream of Wnt/β-catenin signaling. *Development* **140**, 2867–78.

Aldiri, I., Xu, B., Wang, L., Chen, X., Hiler, D., Griffiths, L., Valentine, M., Shirinifard, A., Thiagarajan, S., Sablauer, A., et al. (2017). The Dynamic Epigenetic Landscape of the Retina During Development, Reprogramming, and Tumorigenesis. *Neuron* **94**, 550–568.e10.

Almeida, R. D., Sottile, V., Loose, M., De Sousa, P. a, Johnson, A. D. and Ruzov, A. (2012a). Semi-quantitative immunohistochemical detection of 5-hydroxymethyl-cytosine reveals conservation of its tissue distribution between amphibians and mammals. *Epigenetics* **7**, 137–40.

Almeida, R. D., Loose, M., Sottile, V., Matsa, E., Denning, C., Young, L., Johnson, A. D., Gering, M. and Ruzov, A. (2012b). 5-Hydroxymethyl-Cytosine Enrichment of Non-Committed Cells Is Not a Universal Feature of Vertebrate Development. *Epigenetics* **7**, 383–9.

Almeida, A. D., Boije, H., Chow, R. W., He, J., Tham, J., Suzuki, S. C. and Harris, W. a (2014). Spectrum of Fates: a new approach to the study of the developing zebrafish retina. *Development* **141**, 1971–80.

Appleby, D. W. and Modak, S. P. (1977). DNA degradation in terminally differentiating lens fiber cells from chick embryos. *Proc. Natl. Acad. Sci. U. S. A.* **74**, 5579–83.

Auer, T. O. and Del Bene, F. (2014). CRISPR/Cas9 and TALEN-mediated knock-in approaches in zebrafish. *Methods*.

Auer, T. O., Duroure, K., De Cian, A., Concordet, J.-P. and Del Bene, F. (2014). Highly efficient CRISPR/Cas9-mediated knock-in in zebrafish by homology-independent DNA repair. *Genome Res.* **24**, 142–53.

Bachman, M., Uribe-Lewis, S., Yang, X., Williams, M., Murrell, A. and Balasubramanian, S. (2014). 5-Hydroxymethylcytosine is a predominantly stable DNA modification. *Nat. Chem.* **6**, 1049–1055.

Bassett, E. a and Wallace, V. a (2012). Cell fate determination in the vertebrate retina. *Trends Neurosci.* **35**, 565–73.

Bassnett, S., Wilmarth, P. a and David, L. L. (2009). The membrane proteome of the mouse lens fiber cell. *Mol. Vis.* **15**, 2448–63.

Bassnett, S., Shi, Y. and Vrensen, G. F. J. M. (2011). Biological glass: structural determinants of eye lens transparency. *Philos. Trans. R. Soc. Lond. B. Biol. Sci.* **366**, 1250–64.

Becker, J. R., Chatterjee, S., Robinson, T. Y., Bennett, J. S., Panáková, D., Galindo,

C. L., Zhong, L., Shin, J. T., Coy, S. M., Kelly, A. E., et al. (2014). Differential activation of natriuretic peptide receptors modulates cardiomyocyte proliferation during development. 335–345.

Bedell, V. M., Wang, Y., Campbell, J. M., Poshusta, T. L., Starker, C. G., Krug Ii, R. G., Tan, W., Penheiter, S. G., Ma, A. C., Leung, A. Y. H., et al. (2012a). In vivo genome editing using a high-efficiency TALEN system. *Nature*.

Bedell, V. M., Wang, Y., Campbell, J. M., Poshusta, T. L., Starker, C. G., Krug Ii, R. G., Tan, W., Penheiter, S. G., Ma, A. C., Leung, A. Y. H., et al. (2012b). In vivo genome editing using a high-efficiency TALEN system. *Nature*.

Belliveau, M. J. and Cepko, C. L. (1999). Extrinsic and intrinsic factors control the genesis of amacrine and cone cells in the rat retina. *Development* **126**, 555–566.

Bernardos, R. L., Lentz, S. I., Wolfe, M. S. and Raymond, P. A. (2005). Notch-Delta signaling is required for spatial patterning and Muller glia differentiation in the zebrafish retina. *Dev. Biol.* **278**, 381–395.

Bibliowicz, J., Tittle, R. and Gross, J. (2011). Towards a better understanding of human eye disease: insights from the zebrafish, Danio rerio. *Prog. Mol. Biol.*

Bird, A. (2002). DNA methylation patterns and epigenetic memory DNA methylation patterns and epigenetic memory. *Genes Dev*. 6–21.

Boije, H., Rulands, S., Dudczig, S., Simons, B. D. and Harris, W. A. (2015). The Independent Probabilistic Firing of Transcription Factors: A Paradigm for Clonal Variability in the Zebrafish Retina. *Dev. Cell* **34**, 532–543.

Booth, M. J., Branco, M. R., Ficz, G., Oxley, D., Krueger, F., Reik, W. and Balasubramanian, S. (2012). Quantitative Sequencing of 5-Methylcytosine and 5-Hydroxymethylcytosine at Single-Base Resolution. *Science (80-.).* **8333**,.

Borday, C., Cabochette, P., Parain, K., Mazurier, N., Janssens, S., Tran, H. T., Sekkali, B., Bronchain, O., Vleminckx, K., Locker, M., et al. (2012). Antagonistic cross-regulation between Wnt and Hedgehog signalling pathways controls post-embryonic retinal proliferation. *Development* **139**, 3499–509.

Bröske, A.-M., Vockentanz, L., Kharazi, S., Huska, M. R., Mancini, E., Scheller, M., Kuhl, C., Enns, A., Prinz, M., Jaenisch, R., et al. (2009). DNA methylation protects hematopoietic stem cell multipotency from myeloerythroid restriction. *Nat. Genet.* **41**, 1207–1215.

Brzezinski, J. a. and Reh, T. a. (2015). Photoreceptor cell fate specification in vertebrates. *Development* **142**, 3263–3273.

Bulger, M. and Groudine, M. (2011). Functional and mechanistic diversity of distal transcription enhancers. *Cell* **144**, 327–339.

Cade, L., Reyon, D., Hwang, W. Y., Tsai, S. Q., Patel, S., Khayter, C., Joung, J. K., Sander, J. D., Peterson, R. T. and Yeh, J.-R. J. (2012). Highly efficient generation of heritable zebrafish gene mutations using homo- and heterodimeric TALENs. *Nucleic Acids Res*. 1–10.

Campos, C., Valente, L. M. P. and Fernandes, J. M. O. (2012). Molecular evolution of zebrafish dnmt3 genes and thermal plasticity of their expression during embryonic development. *Gene* **500**, 93–100.

Carlson, D. F., Tan, W., Lillico, S. G., Stverakova, D., Proudfoot, C., Christian, M., Voytas, D. F., Long, C. R., Whitelaw, C. B. a and Fahrenkrug, S. C. (2012). Efficient TALEN-mediated gene knockout in livestock. *Proc. Natl. Acad. Sci. U. S. A*. 1–6.

Carmany-Rampey, A. and Moens, C. B. (2006). Modern mosaic analysis in the zebrafish. *Methods* **39**, 228–238.

Cayouette, M., Barres, B. a and Raff, M. (2003). Importance of intrinsic mechanisms in cell fate decisions in the developing rat retina. *Neuron* **40**, 897–904.

Centanin, L. and Wittbrodt, J. (2014). Retinal neurogenesis. *Development* **141**, 241–4.

Cepko, C. (2014). Intrinsically different retinal progenitor cells produce specific types of progeny. *Nat. Rev. Neurosci*. **15**, 615–627.

Cermak, T., Doyle, E. L., Christian, M., Wang, L., Zhang, Y., Schmidt, C., Baller, J. a, Somia, N. V, Bogdanove, A. J. and Voytas, D. F. (2011). Efficient design and assembly of custom TALEN and other TAL effector-based constructs for DNA targeting. *Nucleic Acids Res*. **39**, e82.

Cerveny, K. L., Cavodeassi, F., Turner, K. J., de Jong-Curtain, T. a, Heath, J. K. and Wilson, S. W. (2010). The zebrafish flotte lotte mutant reveals that the local retinal environment promotes the differentiation of proliferating precursors emerging from their stem cell niche. *Development* **137**, 2107–15.

Chen, B., Dodge, M. E., Tang, W., Lu, J., Ma, Z., Fan, C.-W., Wei, S., Hao, W., Kilgore, J., Williams, N. S., et al. (2009). Small molecule-mediated disruption of Wnt-dependent signaling in tissue regeneration and cancer. *Nat. Chem. Biol*. **5**, 100–7.

Chiodini, F., Matter-Sadzinski, L., Rodrigues, T., Skowronska-Krawczyk, D., Brodier, L., Schaad, O., Bauer, C., Ballivet, M. and Matter, J. M. (2013). A positive feedback loop between ATOH7 and a notch effector regulates cell-cycle progression and neurogenesis in the retina. *Cell Rep*. **3**, 796–807.

Clark, S. J., Harrison, J., Paul, C. L. and Frommer, M. (1994). High sensitivity mapping of methylated cytosines. *Nucleic Acids Res*. **22**, 2990–2997.

Clark, S. J., Argelaguet, R., Kapourani, C.-A., Stubbs, T. M., Lee, H. J., Alda-Catalinas, C., Krueger, F., Sanguinetti, G., Kelsey, G., Marioni, J. C., et al. (2018). scNMT-seq enables joint profiling of chromatin accessibility DNA methylation and transcription in single cells. *Nat. Commun*.

Colquitt, B. M., Allen, W. E., Barnea, G. and Lomvardas, S. (2013). Alteration of genic 5-hydroxymethylcytosine patterning in olfactory neurons correlates with changes in gene expression and cell identity. *PNAS* **110**, 14682–14687.

Cvekl, A. and Duncan, M. K. (2007). Genetic and epigenetic mechanisms of gene regulation during lens development. *Prog. Retin. Eye Res*. **26**, 555–97.

Dahlem, T. J., Hoshijima, K., Jurynec, M. J., Gunther, D., Starker, C. G., Locke, A. S., Weis, A. M., Voytas, D. F. and Grunwald, D. J. (2012). Simple Methods for Generating and Detecting Locus-Specific Mutations Induced with TALENs in the Zebrafish Genome. *PLoS Genet*. **8**, e1002861.

Dai, H.-Q., Wang, B.-A., Yang, L., Chen, J.-J., Zhu, G.-C., Sun, M.-L., Ge, H.,

Wang, R., Chapman, D. L., Tang, F., et al. (2016). TET-mediated DNA demethylation controls gastrulation by regulating Lefty-Nodal signalling. *Nature* **538**, 1–20.

Daron, J. and Slotkin, R. K. (2017). EpiTEome: Simultaneous detection of transposable element insertion sites and their DNA methylation levels. *Genome Biol.* **18**, 1–10.

Das, A. V., James, J., Bhattacharya, S., Imbalzano, A. N., Antony, M. L., Hegde, G., Zhao, X., Mallya, K., Ahmad, F., Knudsen, E., et al. (2007). SWI/SNF chromatin remodeling ATPase Brm regulates the differentiation of early retinal stem cells/progenitors by influencing Brn3b expression and Notch signaling. *J. Biol. Chem.* **282**, 35187–35201.

Davis, A., Matzuk, M. and Reh, T. (2000). Activin A Promotes Progenitor Differentiation into Photoreceptors in Rodent Retina. *Mol. Cell. Neurosci.* **15**, 11–21.

Dawlaty, M. M., Breiling, A., Le, T., Raddatz, G., Barrasa, M. I., Cheng, A. W., Gao, Q., Powell, B. E., Li, Z., Xu, M., et al. (2013). Combined deficiency of Tet1 and Tet2 causes epigenetic abnormalities but is compatible with postnatal development. *Dev. Cell* **24**, 310–23.

Dawlaty, M. M., Breiling, A., Le, T., Barrasa, M. I., Raddatz, G., Gao, Q., Powell, B. E., Cheng, A. W., Faull, K. F., Lyko, F., et al. (2014). Loss of tet enzymes compromises proper differentiation of embryonic stem cells. *Dev. Cell* **29**, 102–11.

Deaton, A. M. and Bird, A. (2011). CpG islands and the regulation of transcription. *Genes Dev.* **25**, 1010–22.

Donato, V. Di, Santis, F. De, Auer, T. O., Testa, N., Sánchez-iranzo, H., Mercader, N., Concordet, J. and Bene, F. Del (2016). 2C-Cas9 : a versatile tool for clonal analysis of gene function. 681–692.

Dovey, H. F., John, V., Anderson, J. P., Chen, L. Z., De Saint Andrieu, P., Fang, L. Y., Freedman, S. B., Folmer, B., Goldbach, E., Holsztynska, E. J., et al. (2001). Functional gamma-secretase inhibitors reduce beta-amyloid peptide levels in brain. *J. Neurochem.* **76**, 173–181.

Dyer, M. A. and Cepko, C. L. (2001). p27 Kip1 and p57 Kip2 Regulate Proliferation in Distinct Retinal Progenitor Cell Populations. *J. neurosc* **21**, 4259–4271.

Ehrlich, M. (2003). Expression of various genes is controlled by DNA methylation during mammalian development. *J. Cell. Biochem.* **88**, 899–910.

Elliott, J., Jolicoeur, C., Ramamurthy, V. and Cayouette, M. (2008). Ikaros Confers Early Temporal Competence to Mouse Retinal Progenitor Cells. *Neuron* **60**, 26–39.

Etchegaray, J.-P., Chavez, L., Huang, Y., Ross, K. N., Choi, J., Martinez-Pastor, B., Walsh, R. M., Sommer, C. A., Lienhard, M., Gladden, A., et al. (2015). The histone deacetylase SIRT6 controls embryonic stem cell fate via TET-mediated production of 5-hydroxymethylcytosine. *Nat. Cell Biol.* **17**, 545–557.

Feng, S., Jacobsen, S. E. and Reik, W. (2010). Epigenetic reprogramming in plant and animal development. *Science* **330**, 622–7.

Furukawa, T., Morrow, E. M. and Cepko, C. L. (1997). Crx, a Novel otx-like Homeobox Gene, Shows Photoreceptor-Specific Expression and Regulates

Photoreceptor Differentiation. *Cell* **91**, 531–541.
Gabher, C. and Wittbrodt, J. (2004). Efficient activation of gene expression using a heat-shock inducible Gal4/Vp16-UAS system in medaka. *BMC Biotechnol.* **4**,.
Gaidatzis, D., Burger, L., Murr, R., Lerch, A., Dessus-Babus, S., Schübeler, D. and Stadler, M. B. (2014). DNA Sequence Explains Seemingly Disordered Methylation Levels in Partially Methylated Domains of Mammalian Genomes. *PLoS Genet.*
Gao, Q., Steine, E. J., Barrasa, M. I., Hockemeyer, D., Pawlak, M., Fu, D., Reddy, S., Bell, G. W. and Jaenisch, R. (2011). Deletion of the de novo DNA methyltransferase Dnmt3a promotes lung tumor progression. *Proc. Natl. Acad. Sci. U. S. A.* **108**, 18061–6.
Garrett-Bakelman, F. E., Sheridan, C. K., Kacmarczyk, T. J., Ishii, J., Betel, D., Alonso, A., Mason, C. E., Figueroa, M. E. and Melnick, A. M. (2015). Enhanced Reduced Representation Bisulfite Sequencing for Assessment of DNA Methylation at Base Pair Resolution. *J. Vis. Exp.*
Ge, L., Zhang, R.-P., Wan, F., Guo, D.-Y., Wang, P., Xiang, L.-X. and Shao, J.-Z. (2014). TET2 Plays an Essential Role in Erythropoiesis by Regulating Lineage-Specific Genes via DNA Oxidative Demethylation in a Zebrafish Model. *Mol. Cell. Biol.* **34**, 989–1002.
Geling, A., Steiner, H., Willem, M., Bally-Cuif, L. and Haass, C. (2002). A γ-secretase inhibitor blocks Notch signaling in vivo and causes a severe neurogenic phenotype in zebrafish. *EMBO Rep.* **3**, 688–694.
Gestri, G., Link, B. A. and Neuhauss, S. C. F. (2012). The visual system of zebrafish and its use to model human ocular Diseases. *Dev. Neurobiol.*
Gifford, C. A., Ziller, M. J., Gu, H., Trapnell, C., Donaghey, J., Tsankov, A., Shalek, A. K., Kelley, D. R., Shishkin, A. A., Issner, R., et al. (2013). Transcriptional and Epigenetic Dynamics during Specification of Human Embryonic Stem Cells. *Cell* **153**, 1149–1163.
Gjini, E., Mansour, M. R., Sander, J. D., Moritz, N., Nguyen, A. T., Kesarsing, M., Gans, E., He, S., Chen, S., Ko, M., et al. (2015). A Zebrafish Model of Myelodysplastic Syndrome Produced through tet2 Genomic Editing. *Mol. Cell. Biol.* **35**, 789–804.
Goll, M. G. and Halpern, M. E. (2011). DNA methylation in zebrafish. *Prog. Mol. Biol. Transl. Sci.* **101**, 193–218.
Gomes, F. L. A. F., Zhang, G., Carbonell, F., Correa, J. A., Harris, W. A., Simons, B. D. and Cayouette, M. (2011). Reconstruction of rat retinal progenitor cell lineages in vitro reveals a surprising degree of stochasticity in cell fate decisions. *Development* **138**, 227–235.
Gore, A. V, Athans, B., Iben, J. R., Johnson, K., Russanova, V., Castranova, D., Pham, V. N., Butler, M. G., Williams-Simons, L., Nichols, J. T., et al. (2016). Epigenetic regulation of hematopoiesis by DNA methylation. *Elife* **5**, e11813.
Greiling, T. M. S. and Clark, J. I. (2009). Early lens development in the zebrafish: a three-dimensional time-lapse analysis. *Dev. Dyn.* **238**, 2254–65.
Greiling, T. M. S., Aose, M. and Clark, J. I. (2010). Cell fate and differentiation of the

developing ocular lens. *Invest. Ophthalmol. Vis. Sci.* **51**, 1540–6.
Hadad, N., Masser, D. R., Logan, S., Wronowski, B., Mangold, C. A., Clark, N., Otalora, L., Unnikrishnan, A., Ford, M. M., Giles, C. B., et al. (2016). Absence of genomic hypomethylation or regulation of cytosine-modifying enzymes with aging in male and female mice. *Epigenetics and Chromatin* **9**, 1–20.
Hahn, M. a, Qiu, R., Wu, X., Li, A. X., Zhang, H., Wang, J., Jui, J., Jin, S.-G., Jiang, Y., Pfeifer, G. P., et al. (2013). Dynamics of 5-hydroxymethylcytosine and chromatin marks in Mammalian neurogenesis. *Cell Rep.* **3**, 291–300.
Hayes, J. M., Hartsock, A., Clark, B. S., Napier, H. R. L., Link, B. a and Gross, J. M. (2012). Integrin α5/Fibronectin1 and Focal adhesion kinase are required for lens fiber morphogenesis in zebrafish. *Mol. Biol. Cell* **23**, 4725–4738.
He, S., Pirity, M. K., Wang, W.-L., Wolf, L., Chauhan, B. K., Cveklova, K., Tamm, E. R., Ashery-Padan, R., Metzger, D., Nakai, A., et al. (2010). Chromatin remodeling enzyme Brg1 is required for mouse lens fiber cell terminal differentiation and its denucleation. *Epigenetics Chromatin* **3**, 21.
He, J., Zhang, G., Almeida, A. D., Cayouette, M., Simons, B. D. and Harris, W. A. (2012). How Variable Clones Build an Invariant Retina. *Neuron* **75**, 786–798.
Heinz, S., Benner, C., Spann, N., Bertolino, E., Lin, Y. C., Laslo, P., Cheng, J. X., Murre, C., Singh, H. and Glass, C. K. (2010). Simple Combinations of Lineage-Determining Transcription Factors Prime cis-Regulatory Elements Required for Macrophage and B Cell Identities. *Mol. Cell* **38**, 576–589.
Hejtmancik, F. (2009). Congenital Cataracts and their Molecular Genetics. **19**, 134–149.
Hino, S., Kishida, S., Michiue, T., Sakamoto, I., Takada, S. and Asashima, M. (2001). Inhibition of the Wnt Signaling Pathway by Idax , a Novel Dvl-Binding Protein Inhibition of the Wnt Signaling Pathway by Idax , a Novel Dvl-Binding Protein. **21**, 330–342.
Horsford, D. J. (2004). Chx10 repression of Mitf is required for the maintenance of mammalian neuroretinal identity. *Development* **132**, 177–187.
Hu, M. and Easter, S. S. (1999). Retinal neurogenesis: the formation of the initial central patch of postmitotic cells. *Dev. Biol.* **207**, 309–21.
Hu, N., Strobl-Mazzulla, P., Sauka-Spengler, T. and Bronner, M. E. (2012). DNA methyltransferase3A as a molecular switch mediating the neural tube-to-neural crest fate transition. *Genes Dev.* **26**, 2380–2385.
Huang, Y., Pastor, W. A., Shen, Y., Tahiliani, M., Liu, D. R. and Rao, A. (2010). The behaviour of 5-hydroxymethylcytosine in bisulfite sequencing. *PLoS One* **5**, 1–9.
Huang, P., Xiao, A., Zhou, M., Zhu, Z., Lin, S. and Zhang, B. (2011). Heritable gene targeting in zebrafish using customized TALENs. *Nat. Biotechnol.* **29**, 699–700.
Huang, Y. H., Su, J., Lei, Y., Brunetti, L., Gundry, M. C., Zhang, X., Jeong, M., Li, W. and Goodell, M. A. (2017). DNA epigenome editing using CRISPR-Cas SunTag-directed DNMT3A. *Genome Biol.* **18**, 1–11.
Hwang, W. Y., Fu, Y., Reyon, D., Maeder, M. L., Tsai, S. Q., Sander, J. D., Peterson, R. T., Yeh, J.-R. J. and Joung, J. K. (2013). Efficient genome editing in zebrafish using a CRISPR-Cas system. *Nat. Biotechnol.* **31**, 227–229.

Illingworth, R. S. and Bird, A. P. (2009). CpG islands--'a rough guide'. *FEBS Lett.* **583**, 1713–20.

Illingworth, R., Kerr, A., Desousa, D., Jørgensen, H., Ellis, P., Stalker, J., Jackson, D., Clee, C., Plumb, R., Rogers, J., et al. (2008). A novel CpG island set identifies tissue-specific methylation at developmental gene loci. *PLoS Biol.* **6**, e22.

Irion, U., Krauss, J. and Nusslein-Volhard, C. (2014). Precise and efficient genome editing in zebrafish using the CRISPR/Cas9 system. *Development* **141**, 4827–30.

Isshiki, T., Pearson, B., Holbrook, S. and Doe, C. Q. (2001). Drosophila neuroblasts sequentially express transcription factors which specify the temporal identity of their neuronal progeny. *Cell.*

Ito, S., D'Alessio, A. C., Taranova, O. V, Hong, K., Sowers, L. C. and Zhang, Y. (2010). Role of Tet proteins in 5mC to 5hmC conversion, ES-cell self-renewal and inner cell mass specification. *Nature* **466**, 1129–33.

Jin, K. (2016). Transitional Progenitors during Vertebrate Retinogenesis. *Mol. Neurobiol.* 1–12.

Jin, B., Ernst, J., Tiedemann, R. L., Xu, H., Sureshchandra, S., Kellis, M., Dalton, S., Liu, C., Choi, J. H. and Robertson, K. D. (2012). Linking DNA Methyltransferases to Epigenetic Marks and Nucleosome Structure Genome-wide in Human Tumor Cells. *Cell Rep.* **2**, 1411–1424.

Johnson, M. B., Wang, P. P., Atabay, K. D., Murphy, E. a, Doan, R. N., Hecht, J. L. and Walsh, C. a (2015). Single-cell analysis reveals transcriptional heterogeneity of neural progenitors in human cortex. *Nat. Neurosci.* 6–8.

Jones, P. A. (2012). Functions of DNA methylation: Islands, start sites, gene bodies and beyond. *Nat. Rev. Genet.* **13**, 484–492.

Jowett, T. and Lettice, L. (1994). Whole-mount in situ hybridizations on zebrafish embryos using a mixture of digoxigenin and fluorescein-labelled probes. **10**, 73–74.

Judson, R. L., Greve, T. S., Parchem, R. J. and Blelloch, R. (2013). MicroRNA-based discovery of barriers to dedifferentiation of fibroblasts to pluripotent stem cells. *Nat. Struct. Mol. Biol.* **20**, 1227–1237.

Jusuf, P. R. and Harris, W. a (2009). Ptf1a is expressed transiently in all types of amacrine cells in the embryonic zebrafish retina. *Neural Dev.* **4**, 34.

Jusuf, P. R., Almeida, A. D., Randlett, O., Joubin, K., Poggi, L. and Harris, W. a (2011). Origin and determination of inhibitory cell lineages in the vertebrate retina. *J. Neurosci.* **31**, 2549–62.

Karlstrom, R. O., Trowe, T., Klostermann, S., Baier, H., Brand, M., Crawford, a D., Grunewald, B., Haffter, P., Hoffmann, H., Meyer, S. U., et al. (1996). Zebrafish mutations affecting retinotectal axon pathfinding. *Development* **123**, 427–38.

Kawakami, K., Abe, G., Asada, T., Asakawa, K., Fukuda, R., Ito, A., Lal, P., Mouri, N., Muto, A., Suster, M. L., et al. (2010). ZTrap: Zebrafish gene trap and enhancer trap database. *BMC Dev. Biol.* **10**,.

Kay, J. N., Finger-Baier, K. C., Roeser, T., Staub, W. and Baier, H. (2001). Retinal Ganglion Cell Genesis Requires lakritz, a Zebrafish atonal Homolog. *Neuron* **30**,

725–736.

Kettleborough, R. N. W., Busch-Nentwich, E. M., Harvey, S. a, Dooley, C. M., de Bruijn, E., van Eeden, F., Sealy, I., White, R. J., Herd, C., Nijman, I. J., et al. (2013). A systematic genome-wide analysis of zebrafish protein-coding gene function. *Nature* **496**, 494–7.

Khoueiry, R., Sohni, A., Thienpont, B., Luo, X., Velde, J. Vande, Bartoccetti, M., Boeckx, B., Zwijsen, A., Rao, A., Lambrechts, D., et al. (2017). Lineage-specific functions of TET1 in the postimplantation mouse embryo. *Nat. Genet.* **49**, 1061–1072.

Kim, C. H., Ueshima, E., Muraoka, O., Tanaka, H., Yeo, S. Y., Huh, T. L. and Miki, N. (1996). Zebrafish elav/HuC homologue as a very early neuronal marker. *Neurosci. Lett.* **216**, 109–112.

Kim, J. H., Lee, S.-R., Li, L.-H., Park, H.-J., Park, J.-H., Lee, K. Y., Kim, M.-K., Shin, B. A. and Choi, S.-Y. (2011). High cleavage efficiency of a 2A peptide derived from porcine teschovirus-1 in human cell lines, zebrafish and mice. *PLoS One* **6**, e18556.

Kim, S. J., Zhao, H., Hardikar, S., Singh, A. K., Goodell, M. A. and Chen, T. (2013a). A DNMT3A mutation common in AML exhibits dominant-negative effects in murine ES cells. *Blood* **122**, 4086–4089.

Kim, M. S., Kim, Y. R., Yoo, N. J. and Lee, S. H. (2013b). Mutational analysis of DNMT3A gene in acute leukemias and common solid cancers. *APMIS* **121**, 85–94.

Kim, R., Sheaffer, K. L., Choi, I., Won, K. and Kaestner, K. H. (2016). Epigenetic regulation of intestinal stem cells by Tet1-mediated DNA hydroxymethylation. 1–10.

Kimmel, C. B., Ballard, W. W., Kimmel, S. R., Ullmann, B. and Schilling, T. F. (1995). Stages of embryonic development of the zebrafish. *Dev. Dyn.* **203**, 253–310.

Klimova, L. and Kozmik, Z. (2014). Stage-dependent requirement of neuroretinal Pax6 for lens and retina development. *Development* **141**, 1292–302.

Ko, M., An, J., Pastor, W. A., Koralov, S. B., Rajewsky, K. and Rao, A. (2015). TET proteins and 5-methylcytosine oxidation in hematological cancers. *Immunol. Rev.* **263**, 6–21.

Kohli, R. M. and Zhang, Y. (2013). TET enzymes, TDG and the dynamics of DNA demethylation. *Nature* **502**, 472–9.

Krueger, F. and Andrews, S. R. (2011). Bismark: A flexible aligner and methylation caller for Bisulfite-Seq applications. *Bioinformatics* **27**, 1571–1572.

Kubo, F., Takeichi, M. and Nakagawa, S. (2005). Wnt2b inhibits differentiation of retinal progenitor cells in the absence of Notch activity by downregulating the expression of proneural genes. *Development* **132**, 2759–2770.

Kubo, N., Toh, H., Shirane, K., Shirakawa, T., Kobayashi, H., Sato, T., Sone, H., Sato, Y., Tomizawa, S., Tsurusaki, Y., et al. (2015). DNA methylation and gene expression dynamics during spermatogonial stem cell differentiation in the early postnatal mouse testis. *BMC Genomics* 1–17.

Kwan, K. M., Fujimoto, E., Grabher, C., Mangum, B. D., Hardy, M. E., Campbell,

D. S., Parant, J. M., Yost, H. J., Kanki, J. P. and Chien, C.-B. (2007). The Tol2kit: a multisite gateway-based construction kit for Tol2 transposon transgenesis constructs. *Dev. Dyn.* **236**, 3088–99.

Kwan, K. M., Otsuna, H., Kidokoro, H., Carney, K. R., Saijoh, Y. and Chien, C.-B. (2012). A complex choreography of cell movements shapes the vertebrate eye. *Development*.

Langley, A. R., Smith, J. C., Stemple, D. L. and Harvey, S. a (2014). New insights into the maternal to zygotic transition. *Development* **141**, 3834–41.

Lee, J., Cox, B. D., Daly, C. M. S., Lee, C., Nuckels, R. J., Tittle, R. K., Uribe, R. a and Gross, J. M. (2012). An ENU mutagenesis screen in zebrafish for visual system mutants identifies a novel splice-acceptor site mutation in patched2 that results in Colobomas. *Invest. Ophthalmol. Vis. Sci.* **53**, 8214–21.

Lee, I., Rasoul, B. A., Holub, A. S., Lejeune, A., Enke, R. A. and Timp, W. (2017). Data Descriptor: Whole genome DNA methylation sequencing of the chicken retina, cornea and brain. *Sci. Data* **4**, 1–8.

Li, E., Beard, C. and Jaenisch, R. (1993). Role for DNA methylation in genomic imprinting. *Nature* **366**, 362–5.

Li, Z., Joseph, N. M. and Easter, S. S. (2000). The morphogenesis of the zebrafish eye, including a fate map of the optic vesicle. *Dev. Dyn.*

Li, H., Handsaker, B., Wysoker, A., Fennell, T., Ruan, J., Homer, N., Marth, G., Abecasis, G. and Durbin, R. (2009). The Sequence Alignment/Map format and SAMtools. *Bioinformatics* **25**, 2078–2079.

Li, T., Yang, D., Li, J., Tang, Y., Yang, J. and Le, W. (2014). Critical Role of Tet3 in Neural Progenitor Cell Maintenance and Terminal Differentiation. *Mol. Neurobiol.*

Li, C., Lan, Y., Schwartz-Orbach, L., Korol, E., Tahiliani, M., Evans, T. and Goll, M. G. (2015). Overlapping Requirements for Tet2 and Tet3 in Normal Development and Hematopoietic Stem Cell Emergence. *Cell Rep.* **12**, 1133–1143.

Li, X., Yue, X., Pastor, W. A., Lin, L., Georges, R., Chavez, L., Evans, S. M. and Rao, A. (2016). Tet proteins influence the balance between neuroectodermal and mesodermal fate choice by inhibiting Wnt signaling. *Proc. Natl. Acad. Sci.* 201617802.

Lio, C.-W. J., Zhang, J., González-Avalos, E., Hogan, P. G., Chang, X. and Rao, A. (2016). Tet2 and Tet3 cooperate with B-lineage transcription factors to regulate DNA modification and chromatin accessibility. *Elife* **5**, e18290.

Liu, W., Khare, S. L., Liang, X., Peters, M. a, Liu, X., Cepko, C. L. and Xiang, M. (2000). All Brn3 genes can promote retinal ganglion cell differentiation in the chick. *Development* **127**, 3237–3247.

Liu, Y., Shen, Y., Rest, J. S., Raymond, P. A. and Zack, D. J. (2001). Isolation and Characterization of a Zebrafish Homologue of the Cone Rod Homeobox Gene. *Investig. Ophthalmol. Vis. Sci.* 481–487.

Lizio, M., Harshbarger, J., Shimoji, H., Severin, J., Kasukawa, T., Sahin, S., Abugessaisa, I., Fukuda, S., Hori, F., Ishikawa-Kato, S., et al. (2015). Gateways to the FANTOM5 promoter level mammalian expression atlas. *Genome Biol.* **16**, 1–

14.
Locker, M., Agathocleous, M., Amato, M. a, Parain, K., Harris, W. a and Perron, M. (2006). controlling the proliferative properties of neural precursors Hedgehog signaling and the retina : insights into the mechanisms controlling the proliferative properties of neural precursors. 3036–3048.
Loh, Y.-H. E., Koemeter-Cox, A., Finelli, M., Shen, L., Friedel, R. H. and Zou, H. (2016). Comprehensive mapping of 5-hydroxymethylcytosine epigenetic dynamics in axon regeneration. *Epigenetics* **0**, 1–16.
Lussi, Y. C., Mariani, L., Friis, C., Peltonen, J., Myers, T. R., Krag, C., Wong, G. and Salcini, A. E. (2016). Impaired removal of H3K4 methylation affects cell fate determination and gene transcription. *Development* dev.139139.
Madzo, J., Liu, H., Rodriguez, A., Vasanthakumar, A., Sundaravel, S., Caces, D. B. D., Looney, T. J., Zhang, L., Lepore, J. B., Macrae, T., et al. (2014). Hydroxymethylation at gene regulatory regions directs stem/early progenitor cell commitment during erythropoiesis. *Cell Rep*. **6**, 231–44.
Marcus, R. C., Delaney, C. L. and Easter, S. S. (1999). Neurogenesis in the visual system of embryonic and adult zebrafish (Danio rerio). *Vis. Neurosci*. **16**, 417–424.
Marquardt, T., Ashery-padan, R., Andrejewski, N., Scardigli, R., Guillemot, F., Gruss, P., Chemistry, M. B., Inserm, C., Pasteur, L. and Ce, I. (2001). Pax6 Is Required for the Multipotent State of Retinal Progenitor Cells. *Cell* **30**, i.
Masai, I., Yamaguchi, M., Tonou-Fujimori, N., Komori, A. and Okamoto, H. (2005). The hedgehog-PKA pathway regulates two distinct steps of the differentiation of retinal ganglion cells: the cell-cycle exit of retinoblasts and their neuronal maturation. *Development* **132**, 1539–53.
Matsushima, D., Heavner, W. and Pevny, L. H. (2011). Combinatorial regulation of optic cup progenitor cell fate by SOX2 and PAX6. *Development* **138**, 443–454.
McKenna, A., Findlay, G. M., Gagnon, J. A., Horwitz, M. S., Schier, A. F. and Shendure, J. (2016). Whole-organism lineage tracing by combinatorial and cumulative genome editing. *Science (80-.).* **353**,.
Meissner, A., Gnirke, A., Bell, G. W., Ramsahoye, B., Lander, E. S. and Jaenisch, R. (2005). Reduced representation bisulfite sequencing for comparative high-resolution DNA methylation analysis. *Nucleic Acids Res*. **33**, 5868–5877.
Meissner, A., Mikkelsen, T. S., Gu, H., Wernig, M., Hanna, J., Sivachenko, A., Zhang, X., Bernstein, B. E., Nusbaum, C., Jaffe, D. B., et al. (2008). Genome-scale DNA methylation maps of pluripotent and differentiated cells. *Nature* **454**, 766–70.
Mellén, M., Ayata, P. and Heintz, N. (2017). 5-Hydroxymethylcytosine Accumulation in Postmitotic Neurons Results in Functional Demethylation of Expressed Genes. *Proc. Natl. Acad. Sci*. 201708044.
Merbs, S. L., Khan, M. a, Hackler, L., Oliver, V. F., Wan, J., Qian, J. and Zack, D. J. (2012). Cell-specific DNA methylation patterns of retina-specific genes. *PLoS One* **7**, e32602.
Meyers, J. R., Hu, L., Moses, A., Kaboli, K., Papandrea, A. and Raymond, P. a

(2012). β-catenin/Wnt signaling controls progenitor fate in the developing and regenerating zebrafish retina. *Neural Dev*. **7**, 30.

Miller, J. C., Tan, S., Qiao, G., Barlow, K. a, Wang, J., Xia, D. F., Meng, X., Paschon, D. E., Leung, E., Hinkley, S. J., et al. (2011). A TALE nuclease architecture for efficient genome editing. *Nat. Biotechnol*. **29**, 143–8.

Mo, A., Mukamel, E. A., Davis, F. P., Luo, C., Henry, G. L., Picard, S., Urich, M. A., Nery, J. R., Sejnowski, T. J., Lister, R., et al. (2015). Epigenomic Signatures of Neuronal Diversity in the Mammalian Brain. *Neuron* **86**, 1369–1384.

Mo, A., Luo, C., Davis, F. P., Mukamel, E. A., Henry, G. L., Nery, J. R., Urich, M. A., Picard, S., Lister, R., Eddy, S. R., et al. (2016). Epigenomic landscapes of retinal rods and cones. *Elife* **5**, 1–29.

Mooijman, D., Dey, S. S., Boisset, J. C., Crosetto, N. and Van Oudenaarden, A. (2016). Single-cell 5hmC sequencing reveals chromosome-wide cell-to-cell variability and enables lineage reconstruction. *Nat. Biotechnol*.

Neff, K. L., Argue, D. P., Ma, A. C., Lee, H. B., Clark, K. J. and Ekker, S. C. (2013). Mojo Hand, a TALEN design tool for genome editing applications. *BMC Bioinformatics* **14**, 1.

Neri, F., Incarnato, D., Krepelova, A., Dettori, D., Rapelli, S., Maldotti, M., Parlato, C., Anselmi, F., Galvagni, F. and Oliviero, S. (2015). TET1 is controlled by pluripotency-associated factors in ESCs and downmodulated by PRC2 in differentiated cells and tissues. *Nucleic Acids Res*. **43**, 6814–6826.

Nishimoto, S., Kawane, K., Watanabe-Fukunaga, R., Fukuyama, H., Ohsawa, Y., Uchiyama, Y., Hashida, N., Ohguro, N., Tano, Y., Morimoto, T., et al. (2003). Nuclear cataract caused by a lack of DNA degradation in the mouse eye lens. *Nature* **424**, 1071–4.

Nishiya, N., Oku, Y., Kumagai, Y., Sato, Y., Yamaguchi, E., Sasaki, A., Shoji, M., Ohnishi, Y., Okamoto, H. and Uehara, Y. (2014). A zebrafish chemical suppressor screening identifies small molecule inhibitors of the Wnt/beta-catenin pathway. *Chem. Biol*. **21**, 530–540.

Obholzer, N., Swinburne, I. A., Schwab, E., Nechiporuk, A. V., Nicolson, T. and Megason, S. G. (2012). Rapid positional cloning of zebrafish mutations by linkage and homozygosity mapping using whole-genome sequencing. *Development* **139**, 4280–4290.

Okae, H., Chiba, H., Hiura, H., Hamada, H. and Sato, A. (2014). Genome-Wide Analysis of DNA Methylation Dynamics during Early Human Development. *PLoS Genet*. **10**, 1–12.

Okano, M., Bell, D. W., Haber, D. a and Li, E. (1999). DNA methyltransferases Dnmt3a and Dnmt3b are essential for de novo methylation and mammalian development. *Cell* **99**, 247–57.

Orlanski, S., Labi, V., Reizel, Y., Spiro, A., Lichtenstein, M., Levin-Klein, R., Koralov, S. B., Skversky, Y., Rajewsky, K., Cedar, H., et al. (2016). Tissue-specific DNA demethylation is required for proper B-cell differentiation and function. *Proc. Natl. Acad. Sci. U. S. A*. **113**, 5018–23.

Oron-Karni, V., Farhy, C., Elgart, M., Marquardt, T., Remizova, L., Yaron, O., Xie, Q., Cvekl, A. and Ashery-Padan, R. (2008). Dual requirement for Pax6 in retinal progenitor cells. *Development* **135**, 4037–4047.

Pan, L., Deng, M., Xie, X. and Gan, L. (2008). ISL1 and BRN3B co-regulate the differentiation of murine retinal ganglion cells. *Development* **135**, 1981–1990.

Pastor, W. a., Aravind, L. and Rao, A. (2013). TETonic shift: biological roles of TET proteins in DNA demethylation and transcription. *Nat. Rev. Mol. Cell Biol.* **14**, 341–356.

Patel, A. and McFarlane, S. (2000). Overexpression of FGF-2 Alters Cell Fate Specification in the Developing Retina of Xenopus laevis. *Dev. Biol.* **222**, 170–180.

Pedersen, B. S., Eyring, K., De, S., Yang, I. V. and Schwartz, D. A. (2014). Fast and accurate alignment of long bisulfite-seq reads. **0**, 1–2.

Perera, A., Eisen, D., Wagner, M., Laube, S. K., K??nzel, A. F., Koch, S., Steinbacher, J., Schulze, E., Splith, V., Mittermeier, N., et al. (2015). TET3 is recruited by REST for context-specific hydroxymethylation and induction of gene expression. *Cell Rep.* **11**, 283–294.

Perron, M. and Harris, W. a (2000). Determination of vertebrate retinal progenitor cell fate by the Notch pathway and basic helix-loop-helix transcription factors. *Cell. Mol. Life Sci.* **57**, 215–223.

Pittman, A. J., Law, M.-Y. and Chien, C.-B. (2008). Pathfinding in a large vertebrate axon tract: isotypic interactions guide retinotectal axons at multiple choice points. *Development* **135**, 2865–71.

Poggi, L., Vitorino, M., Masai, I. and Harris, W. A. (2005). Influences on neural lineage and mode of division in the zebrafish retina in vivo. *J. Cell Biol.*

Qu, J., Zhou, M., Song, Q., Hong, E. E. and Smith, A. D. (2013). MLML: Consistent simultaneous estimates of DNA methylation and hydroxymethylation. *Bioinformatics* **29**, 2645–2646.

Quastler, H. and Sherman, F. G. (1959). Cell population kinetics in the intestinal epithelium of the mouse. *Exp. Cell Res.* **17**, 420–438.

Quinlan, A. R. and Hall, I. M. (2010). BEDTools: A flexible suite of utilities for comparing genomic features. *Bioinformatics* **26**, 841–842.

Rai, K., Jafri, I. F., Chidester, S., James, S. R., Karpf, A. R., Cairns, B. R. and Jones, D. a (2010). Dnmt3 and G9a cooperate for tissue-specific development in zebrafish. *J. Biol. Chem.* **285**, 4110–21.

Raiber, E.-A., Beraldi, D., Martínez Cuesta, S., McInroy, G. R., Kingsbury, Z., Becq, J., James, T., Lopes, M., Allinson, K., Field, S., et al. (2017). Base resolution maps reveal the importance of 5-hydroxymethylcytosine in a human glioblastoma. *npj Genomic Med.* **2**, 6.

Rapaport, D. H., Patheal, S. L. and Harris, W. A. (2001). Cellular competence plays a role in photoreceptor differentiation in the developing Xenopus retina. *J. Neurobiol.*

Raymond, P. a, Barthel, L. K., Bernardos, R. L. and Perkowski, J. J. (2006). Molecular characterization of retinal stem cells and their niches in adult zebrafish. *BMC Dev. Biol.* **6**, 36.

Ren, B. and Pott, S. (2017). Simultaneous measurement of chromatin accessibility, DNA methylation, and nucleosome phasing in single cells.

Rhee, K.-D., Yu, J., Zhao, C. Y., Fan, G. and Yang, X.-J. (2012). Dnmt1-dependent DNA methylation is essential for photoreceptor terminal differentiation and retinal neuron survival. *Cell Death Dis.* **3**, e427.

Rieger, M. A. and Schroeder, T. (2012). Hematopoiesis. *Cold Spring Harb. Perspect. Biol.*

Robertson, A. B., Dahl, J. A., Vgbø, C. B., Tripathi, P., Krokan, H. E. and Klungland, A. (2011). A novel method for the efficient and selective identification of 5-hydroxymethylcytosine in genomic DNA. *Nucleic Acids Res.* **39**,.

Russler-Germain, D. A., Spencer, D. H., Young, M. A., Lamprecht, T. L., Miller, C. A., Fulton, R., Meyer, M. R., Erdmann-Gilmore, P., Townsend, R. R., Wilson, R. K., et al. (2014). The R882H DNMT3A Mutation Associated with AML Dominantly Inhibits Wild-Type DNMT3A by Blocking Its Ability to Form Active Tetramers. *Cancer Cell* **25**, 442–454.

Sakuma, T., Hosoi, S., Woltjen, K., Suzuki, K.-I., Kashiwagi, K., Wada, H., Ochiai, H., Miyamoto, T., Kawai, N., Sasakura, Y., et al. (2013). Efficient TALEN construction and evaluation methods for human cell and animal applications. *Genes Cells* **18**, 315–26.

Santana, A. and Waiswol, M. (2011). The genetic and molecular basis of congenital cataract. **74**, 136–142.

Santiago, M., Antunes, C., Guedes, M., Sousa, N. and Marques, C. J. (2014). TET enzymes and DNA hydroxymethylation in neural development and function - How critical are they? *Genomics* **104**, 334–340.

Schmitt, E. A. and Dowling, J. E. (1994). Early Eye Morphogenesis in the Zebrafish, Brachydanio rerio. *J. Comp. Neurol.* 344–532.

Schmitt, E. a and Dowling, J. E. (1996). Comparison of topographical patterns of ganglion and photoreceptor cell differentiation in the retina of the zebrafish, Danio rerio. *J. Comp. Neurol.* **371**, 222–34.

Schmitt, E. A. and Dowling, J. E. (1999). Early Retinal Development in the Zebrafish , Danio rerio : Light and Electron Microscopic Analyses. **536**, 515–536.

Schroeder, D. I., Blair, J. D., Lott, P., Yu, H. O. K., Hong, D., Crary, F., Ashwood, P., Walker, C., Korf, I., Robinson, W. P., et al. (2013). The human placenta methylome. *Proc. Natl. Acad. Sci.* **110**, 6037–6042.

Sen, G. L., Reuter, J. a, Webster, D. E., Zhu, L. and Khavari, P. a (2010). DNMT1 maintains progenitor function in self-renewing somatic tissue. *Nature* **463**, 563–7.

Seritrakul, P. and Gross, J. M. (2014). Expression of the de novo DNA methyltransferases (dnmt3 - dnmt8) during zebrafish lens development. *Dev. Dyn.* **243**, 350–6.

Seritrakul, P. and Gross, J. M. (2017). *Tet-mediated DNA hydroxymethylation regulates retinal neurogenesis by modulating cell-extrinsic signaling pathways*.

Shen, L., Wu, H., Diep, D., Yamaguchi, S., D'Alessio, A. C., Fung, H. L., Zhang, K. and Zhang, Y. (2013). Genome-wide analysis reveals TET- and TDG-dependent 5-

methylcytosine oxidation dynamics. *Cell* **153**, 692–706.

Shimoda, N., Yamakoshi, K., Miyake, A. and Takeda, H. (2005). Identification of a gene required for de novo DNA methylation of the zebrafish no tail gene. *Dev. Dyn.* **233**, 1509–16.

Singh, R. K., Mallela, R. K., Hayes, A., Dunham, N. R., Hedden, M. E., Enke, R. A., Fariss, R. N., Sternberg, H., West, M. D. and Nasonkin, I. O. (2016). Dnmt1, Dnmt3a and Dnmt3b cooperate in photoreceptor and outer plexiform layer development in the mammalian retina. *Exp. Eye Res.*

Slater, J. L., Landman, K. A., Hughes, B. D., Shen, Q. and Temple, S. (2009). Cell lineage tree models of neurogenesis. *J. Theor. Biol.*

Slotkin, R. K. and Martienssen, R. (2007). Transposable elements and the epigenetic regulation of the genome. *Nat. Rev. Genet.* **8**, 272–285.

Smith, Z. D. and Meissner, A. (2013). DNA methylation: roles in mammalian development. *Nat. Rev. Genet.* **14**, 204–20.

Song, Q., Garvin, T., Smith, A. and Qu, J. (2014). The Smithlab DNA Methylation Data Analysis Pipeline (MethPipe) Methylome construction Mapping reads. 1–19.

Sullivan, C. H. and Grainger, R. M. (1986). Delta-crystallin genes become hypomethylated in postmitotic lens cells during chicken development. *PNAS* **83**, 329–333.

Sullivan, C. H., Norman, J. T., Borrás, T. and Grainger, R. M. (1989). Developmental regulation of hypomethylation of delta-crystallin genes in chicken embryo lens cells. *Mol. Cell. Biol.* **9**, 3132–5.

Suzuki, M. M. and Bird, A. (2008). DNA methylation landscapes: provocative insights from epigenomics. *Nat. Rev. Genet.* **9**, 465–76.

Szwagierczak, A., Bultmann, S., Schmidt, C. S., Spada, F. and Leonhardt, H. (2010). Sensitive enzymatic quantification of 5-hydroxymethylcytosine in genomic DNA. *Nucleic Acids Res.* **38**,.

Tahiliani, M., Koh, K. P., Shen, Y., Pastor, W. a, Bandukwala, H., Brudno, Y., Agarwal, S., Iyer, L. M., Liu, D. R., Aravind, L., et al. (2009). Conversion of 5-methylcytosine to 5-hydroxymethylcytosine in mammalian DNA by MLL partner TET1. *Science* **324**, 930–5.

Thisse, B. and Thisse, C. (2004). Fast Release Clones : A High Throughput Expression Analysis. *ZFIN Direct Data Submiss.*

Tittle, R. K., Sze, R., Ng, A., Nuckels, R. J., Swartz, M. E., Anderson, R. M., Bosch, J., Stainier, D. Y. R., Eberhart, J. K. and Gross, J. M. (2011). Uhrf1 and Dnmt1 are required for development and maintenance of the zebrafish lens. *Dev. Biol.* **350**, 50–63.

Trapnell, C., Williams, B. a, Pertea, G., Mortazavi, A., Kwan, G., van Baren, M. J., Salzberg, S. L., Wold, B. J. and Pachter, L. (2010). Transcript assembly and quantification by RNA-Seq reveals unannotated transcripts and isoform switching during cell differentiation. *Nat. Biotechnol.* **28**, 511–515.

Trapnell, C., Roberts, A., Goff, L., Pertea, G., Kim, D., Kelley, D. R., Pimentel, H., Salzberg, S. L., Rinn, J. L. and Pachter, L. (2012). Differential gene and

transcript expression analysis of RNA-seq experiments with TopHat and Cufflinks. *Nat. Protoc.* **7**, 562–78.

Trevarrow, B. and Kimmel, B. (1990). Organization of Hindbrain Segments in the Zebrafish Embryo. *Neuron* **4**, 669–679.

Trimarchi, J. M., Stadler, M. B. and Cepko, C. L. (2008). Individual retinal progenitor cells display extensive heterogeneity of gene expression. *PLoS One* **3**,.

Trowbridge, J. J., Snow, J. W., Kim, J. and Orkin, S. H. (2009). DNA Methyltransferase 1 Is Essential for and Uniquely Regulates Hematopoietic Stem and Progenitor Cells. *Cell Stem Cell* **5**, 442–449.

Ueno, K., Iwagawa, T., Kuribayashi, H., Baba, Y., Nakauchi, H., Murakami, A., Nagasaki, M., Suzuki, Y. and Watanabe, S. (2016). Transition of differential histone H3 methylation in photoreceptors and other retinal cells during retinal differentiation. *Sci. Rep.* **6**, 29264.

Uribe, R. a and Gross, J. M. (2007). Immunohistochemistry on cryosections from embryonic and adult zebrafish eyes. *CSH Protoc.* **2007**, pdb.prot4779.

Uribe, R. a and Gross, J. M. (2010). Id2a influences neuron and glia formation in the zebrafish retina by modulating retinoblast cell cycle kinetics. *Development* **137**, 3763–74.

Uribe, R. a, Kwon, T., Marcotte, E. M. and Gross, J. M. (2012). Id2a functions to limit Notch pathway activity and thereby influence the transition from proliferation to differentiation of retinoblasts during zebrafish retinogenesis. *Dev. Biol.* **371**, 280–92.

VanderKraats, N. D., Hiken, J. F., Decker, K. F. and Edwards, J. R. (2013). Discovering high-resolution patterns of differential DNA methylation that correlate with gene expression changes. *Nucleic Acids Res.* 1–12.

Vitorino, M., Jusuf, P. R., Maurus, D., Kimura, Y., Higashijima, S.-I. and Harris, W. a (2009). Vsx2 in the zebrafish retina: restricted lineages through derepression. *Neural Dev.* **4**, 14.

Wallner, S., Schröder, C., Leitão, E., Berulava, T., Haak, C., Beißer, D., Rahmann, S., Richter, A. S., Manke, T., Bönisch, U., et al. (2016). Epigenetic dynamics of monocyte-to-macrophage differentiation. *Epigenetics and Chromatin* **9**, 1–17.

Walls, G. L. (1963). The vertebrate eye and its adaptive radiation. *Optom. Vis. Sci.* 693–719.

Walsh, C. P., Chaillet, R. and Bestor, T. H. (1998). Transcription of IAP endogenous retroviruses is constrained by cytosine methylation. *Nat. Genet.* **20**, 116–117.

Wang, X., Emelyanov, A., Korzh, V. and Gong, Z. (2003). Zebrafish atonal homologue zath3 is expressed during neurogenesis in embryonic development. *Dev. Dyn.* **227**, 587–592.

Wang, L., Hiler, D., Xu, B., AlDiri, I., Chen, X., Zhou, X., Griffiths, L., Valentine, M., Shirinifard, A., Sablauer, A., et al. (2018). Retinal Cell Type DNA Methylation and Histone Modifications Predict Reprogramming Efficiency and Retinogenesis in 3D Organoid Cultures. *Cell Rep.* **22**, 2601–2614.

Watanabe, T. and Raff, M. C. (1990). Rod photoreceptor development in vitro:

Intrinsic properties of proliferating neuroepithelial cells change as development proceeds in the rat retina. *Neuron* **4**, 461–467.

Wetts, R. and Fraser, S. (1988). Multipotent precursors can give rise to all major cell types of the frog retina. *Science (80-.).* **239**, 1142–1145.

Wienholds, E. (2005). MicroRNA Expression in Zebrafish Embryonic Development. *Science (80-.).* **309**, 310–311.

Wong, L. L. and Rapaport, D. H. (2009). Defining retinal progenitor cell competence in Xenopus laevis by clonal analysis. *Development*.

Woo, K. and Fraser, S. E. (1995). Order and coherence in the fate map of the zebrafish nervous system. *Development* **121**, 2595–609.

Wride, M. a (2011). Lens fibre cell differentiation and organelle loss: many paths lead to clarity. *Philos. Trans. R. Soc. Lond. B. Biol. Sci.* **366**, 1219–33.

Wu, H. and Zhang, Y. (2011). Mechanisms and functions of Tet protein- mediated 5-methylcytosine oxidation. *Genes Dev.* **25**, 2436–2452.

Wu, H. and Zhang, Y. (2014). Reversing DNA Methylation: Mechanisms, Genomics, and Biological Functions. *Cell* **156**, 45–68.

Wu, H., Coskun, V., Tao, J., Xie, W., Ge, W., Yoshikawa, K., Li, E., Zhang, Y. and Sun, Y. E. (2010). Dnmt3a-dependent nonpromoter DNA methylation facilitates transcription of neurogenic genes. *Science* **329**, 444–8.

Wu, Z., Huang, K., Yu, J., Le, T., Namihira, M., Liu, Y., Zhang, J., Xue, Z., Cheng, L. and Fan, G. (2012). Dnmt3a regulates both proliferation and differentiation of mouse neural stem cells. *J. Neurosci. Res.* **90**, 1883–91.

Xu, Y., Wu, F., Tan, L., Kong, L., Xiong, L., Deng, J., Barbera, A. J., Zheng, L., Zhang, H., Huang, S., et al. (2011). Genome-wide regulation of 5hmC, 5mC, and gene expression by Tet1 hydroxylase in mouse embryonic stem cells. *Mol. Cell* **42**, 451–64.

Xu, Y., Xu, C., Kato, A., Tempel, W., Abreu, J. G., Bian, C., Hu, Y., Hu, D., Zhao, B., Cerovina, T., et al. (2012). Tet3 CXXC domain and dioxygenase activity cooperatively regulate key genes for Xenopus eye and neural development. *Cell* **151**, 1200–13.

Yamaguchi, M., Tonou-Fujimori, N., Komori, A., Maeda, R., Nojima, Y., Li, H., Okamoto, H. and Masai, I. (2005). Histone deacetylase 1 regulates retinal neurogenesis in zebrafish by suppressing Wnt and Notch signaling pathways. *Development* **132**, 3027–43.

Yamaguchi, S., Hong, K., Liu, R., Shen, L., Inoue, A., Diep, D., Zhang, K. and Zhang, Y. (2012). Tet1 controls meiosis by regulating meiotic gene expression. *Nature* **492**, 443–7.

Yang, X.-J. (2004). Roles of cell-extrinsic growth factors in vertebrate eye pattern formation and retinogenesis. *Semin. Cell Dev. Biol.* **15**, 91–103.

Yaron, O., Farhy, C., Marquardt, T., Applebury, M. and Ashery-Padan, R. (2006). Notch1 functions to suppress cone-photoreceptor fate specification in the developing mouse retina. *Development* **133**, 1367–1378.

Yu, M., Hon, G. C., Szulwach, K. E., Song, C.-X., Jin, P., Ren, B. and He, C. (2012).

Tet-assisted bisulfite sequencing of 5-hydroxymethylcytosine. *Nat. Protoc.* **7**, 2159–70.

Zemach, A., McDaniel, I. E., Silva, P. and Zilberman, D. (2010). Genome-wide evolutionary analysis of eukaryotic DNA methylation. *Science* **328**, 916–9.

Zentner, G. E. and Scacheri, P. C. (2012). The chromatin fingerprint of gene enhancer elements. *J. Biol. Chem.* **287**, 30888–30896.

Zhou, P., Luo, Y., Liu, X., Fan, L. and Lu, Y. (2012). Down-regulation and CpG island hypermethylation of CRYAA in age-related nuclear cataract. *FASEB J.* **26**, 4897–902.

Ziller, M. J., Gu, H., Müller, F., Donaghey, J., Tsai, L. T.-Y., Kohlbacher, O., De Jager, P. L., Rosen, E. D., Bennett, D. a, Bernstein, B. E., et al. (2013). Charting a dynamic DNA methylation landscape of the human genome. *Nature* **500**, 477–81.

Vita

Pawat Seritrakul was born in Bangkok, Thailand in 1987, the son of Suvit Seritrakul and Wilailuk Seritrakul. After graduating from Mahidol Wittayanusorn in Nakhon Pathom province in Thailand and Blair Academy in New Jersey in 2007, he attended Bowdoin College in Brunswick, Maine, receiving a Bachelor of Arts (*magna cum laude*) with a major in Biology and a minor in Visual Arts in May, 2011. In September 2011, he enrolled in the Institute for Cellular and Molecular Biology graduate program in the Graduate School of the University of Texas at Austin.

E-mail: pawat@utexas.edu

This dissertation was typed by the author.